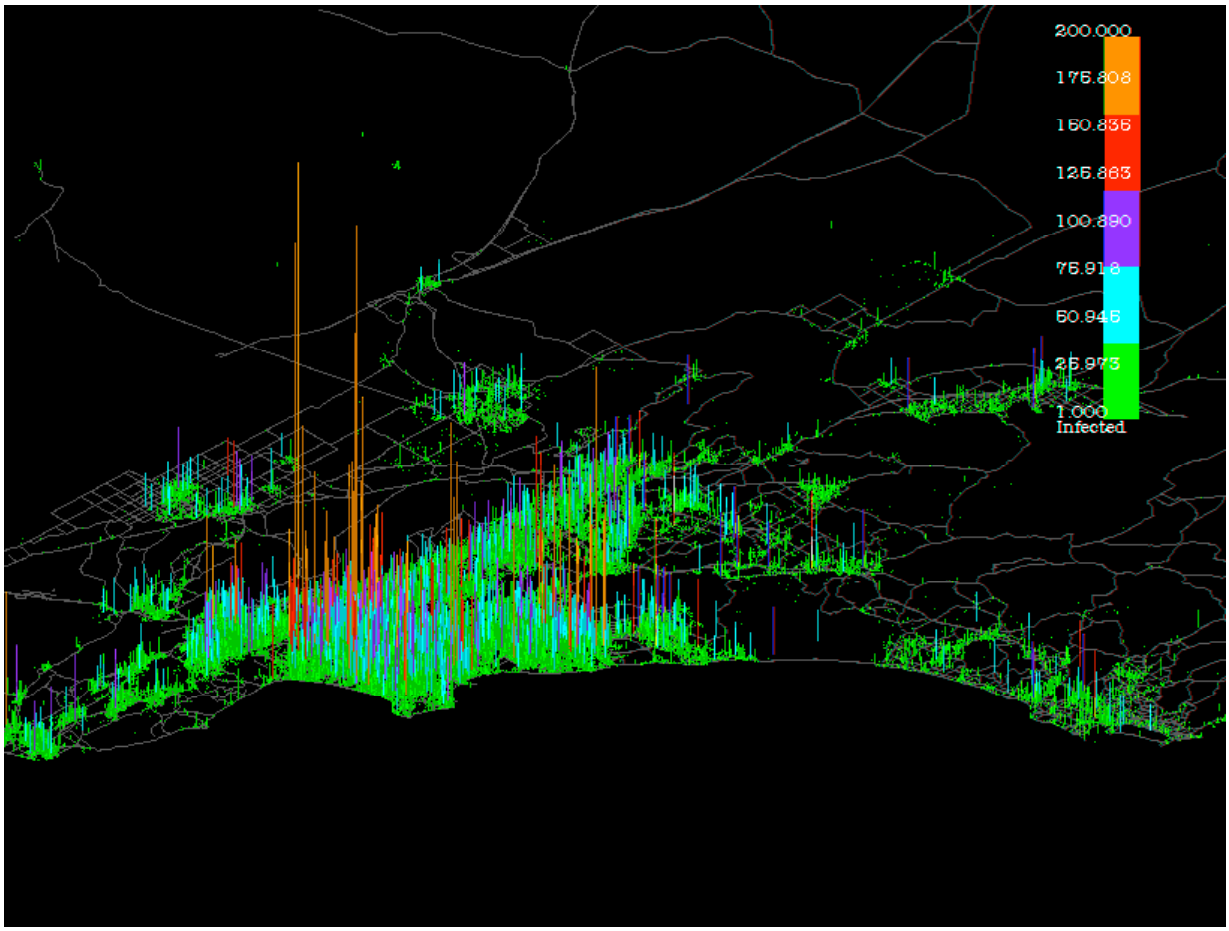
EpiSimS Los Angeles Case Study

Los Alamos National Laboratory
LAUR-06-0666
January 31, 2006

Study participants: S.Y. Del Valle, D. Kubicek, S.M. Mniszewski, J.M. Riese, P.R. Romero, J.P. Smith, P.D. Stroud, S.J. Sydoriak.

Edited by Phillip D. Stroud, LANL, 505-667-6654, stroud@lanl.gov
Sponsoring Agency: DHS-IP. Point of Contact: Jon MacLaren, 202-282-8719, jon.m.maclaren@dhs.gov
NISAC Point of Contact: Randy E. Michelsen, LANL, 505-665-1522, rem@lanl.gov
EpiSimS Project Point of Contact: James P. Smith, LANL, 505-665-0921, jpsmith@lanl.gov



Executive Summary

The Los Angeles Case Study is a milestone in the development of EpiSimS (Epidemic Simulation System). Several new capabilities were implemented into EpiSimS, to advance the simulation capability, to enable high-fidelity simulation of an influenza pandemic, and to model realistic epidemic response policies.

A synthetic population has been built to match the 2000 US Census data, representing over 16 million persons in the five counties of Los Angeles, Orange, Riverside, Ventura, and San Bernardino as individual agents. This synthetic population captures demographic and geographic data down to the census-block or finer resolution. The geographic representation of the region is captured in 562,452 locations, representing households, small neighborhoods, workplaces, schools, etc. Each individual has a set of activities, based on demographics. A schedule is generated for each individual, specifying where and when they travel during the day. Persons that occupy the same room at the same time have opportunity to transmit disease. The social contact structure, including number of contacts per person, and the strength, duration, and type of such contacts, thus emerges from the simulation. Because there is no way to observe or quantify the actual social contact structure in Los Angeles, the social contact structure that emerges from EpiSimS must be regarded as the best such information that exists.

The first step in the process of creating the synthetic population is conducted with a set of software utilities called UPMoST. The UPMoST team gathered population, household and activity data using US Census Bureau Data, the Dun and Bradstreet Business Database, the National Household Travel Survey, the National Center of Educational Statistics, and NGA (National Geospatial-Intelligence Agency) Emergency Response Services. UPMoST person and household information is used to generate demographic data for each person in the simulation, including each person's id, age, gender, economic status, family situation, whether they are a worker or not, and home location.

The next step is to generate activity schedules for each individual. The National Household Travel Survey is used to assign a set of activities to each individual, to specify the order in which they are undertaken during the day, and how long is spent in each activity. Each individual is assigned an activity schedule depending on their demographic data and their role within their household.

The third step determines the location where each activity of each individual takes place. This assignment is performed by an algorithm that accounts for the home address of the individual, the spatial distribution of all locations at which the activity occurs, and the number of people participating in the activity at each location (e.g. the business directory data specifies the number of employees of various job categories by address).

The fourth step, partitioning, divides each location into the appropriate number of mixing rooms, or sub-locations, for each activity type. Thus a school location is partitioned into classrooms, work locations are partitioned into mixing groups, shopping malls are partitioned into stores, etc. Finally, a small subset of the population is selected to represent the initial infected index cases.

A model of the disease progression of pandemic influenza has been build and implemented into EpiSimS. The model implements the details of progression thorough incubation stage, with data-based transition times to infectiousness and onset of symptoms. The duration of the infectious stage is also drawn from data-based histograms. A subclinical disease manifestation has been implemented, and accounts for one third of all infections. Persons with subclinical manifestations continue with their normal pattern of activity, and are contagious although at a lower rate than symptomatic persons. The disease model accounts for behavior modification in infected persons: a symptomatic person may forego his normal activities and instead remain at home. The disease model is intended to allow for examination of alternative strategies of treating pneumonia complications, but this capability has not been pursued in the Los Angeles case study. A fatality rate of 2% of symptomatic cases regardless of age is used for consistency with the 1918 Spanish flu pandemic. The effectiveness of available antiviral medications and vaccines are represented at a moderately detailed level. Scenarios involving response with these treatments have been constructed to be consistent with current stockpiles, and near-term projected stockpiles.

Historical stage duration distributions for incubation and infectiousness of influenza have been implemented into EpiSimS as half-day histograms, based on survey literature. A basic disease manifestation has been developed, having four primary disease states (incubating, subclinical, symptomatic circulating, and symptomatic non-circulating), plus three additional states (susceptible, recovered, and dead). The transition parameters, stage duration histograms, infectiousness depend on the age, and on whether, when, and what treatment has been given. For convenience, four age groups are implemented as different disease manifestations. Also for convenience, the four primary disease states have been expanded to thirteen to explicitly represent the treatment states.

The base case scenario specifies a target attack rate (i.e. the fraction of the population that gets infected during the epidemic) of 25% in the absence of effective treatments. Since the number and strength of contacts are unknown prior to running the simulation, the infectiousness had to be calibrated to produce the target attack rate. A scoping model, epiHist, was developed to project the epidemic curve, obtained by an EpiSimS simulation of a few days of the epidemic, to completion. The calibration process found a basic adult-to-adult infectiousness value of 0.00285 transmission probability per hour per infectious person per susceptible contact. This value is about 60% lower than our original estimates based on the literature. This basic infectiousness applies for household and work contacts, and a reduced infectiousness applies to casual contacts such as might occur while shopping.

Several computational advances have been implemented into EpiSimS, resulting in significant gains in efficiency. A new algorithm has reduced the time required to initially distribute individuals and locations to processors from 4 hours down to only 45 minutes, for the Los Angeles scenario. By aggregating locations into links, the simulation can now assign persons to sub-locations (i.e. mixing rooms) more efficiently. Several new scripts rapidly extract select data from the large EpiSimS output event files. EpiSimS has been successfully ported to a 64-bit architecture cluster.

Two main sets of results are reported in the Los Angeles Case Study. The first are for an influenza pandemic in which no effective vaccine or antiviral treatments are available. The second examines the impact of various outbreak response strategies.

The base case influenza pandemic (with no effective antiviral or vaccine) shows an early growth rate of 7.2%, corresponding to a reproductive number of 1.34. The epidemic grows from 202 index cases to a peak of 65,278 new infections per day on day 125. The peak incidence rate is attained on day 129, when 2.41% of the population is in some stage of infection. At the peak, 1.1% of the population is symptomatic. About 25% of the population gets infected, including subclinical cases.

More infections (44%) are acquired at home than anywhere else, followed by work (29%) and school (19%). The epidemic displays an early wave of school-acquired infections.

During the Los Angeles case study calibration process, it was discovered that the expected number of transmissions per infectious person follows a power-law dependence on the susceptible fraction of the population. This potentially important observation results in a much smaller attack rate for a given reproductive number, relative to traditional epidemic modeling with a homogeneous mixing assumption. An article on this discovery has been accepted for publication in Mathematical Biosciences.

Some visualization products have been created to show the spatial distribution of infection through the course of the epidemic. For the Los Angeles Case Study, the epidemic curve as a fraction of the population is uniform across the five counties. There is some interesting spatial structure, but some further tools and analysis are required for statistical analysis of the spatial dynamics of epidemics.

The impact of antiviral medication has been examined in a series of EpiSimS simulations. The antiviral scenario assumes that a fraction of symptomatic persons are diagnosed, make a list of their contacts, and that a fraction of these contacts are traced and treated. The implementation of contact tracing also would apply for a contact-tracing & quarantine strategy. The resulting epidemic curve and total number of infections is stochastic, and depends on the fraction of contacts that get named, and the fraction of named contacts that get traced, and on when the contact tracing response is initiated. For reasonable naming and tracing fractions (all household members, most classmates and coworkers), even partially effective antivirals are found to effectively prevent an epidemic.

For mass vaccination, assuming roughly the same vaccine effectiveness as in normal flu seasons, EpiSimS simulations show that vaccination of ~40% of the population (as occurs during most flu seasons in the US) will effectively prevent an epidemic. Vaccination of ~20% of the population (as occurred during the 2004-2005 flu season), whether given at random or target to children and seniors, reduces the size of the epidemic, but the epidemic is still substantial.

Contents

1 INTRODUCTION 6

2 EPISIMS SIMULATION SET-UP 6

2.1 CONSTRUCTION OF THE SYNTHETIC LOS ANGELES POPULATION 6
2.2 SCENARIO GENERATION 12

3 PANDEMIC INFLUENZA DISEASE MODELING 16

3.1 INFLUENZA DISEASE PROGRESSION 16
3.2 VACCINATION AND ANTIVIRAL MEDICATION 22
3.3 EPISIMS INFECTIOUSNESS CALIBRATION 26

4 EPISIMS NEW CAPABILITIES 28

4.1 DISEASE PROGRESSION MODEL ENHANCEMENTS 28
4.2 IMPLEMENTATION OF CONSEQUENCE MITIGATION STRATEGIES 32
4.3 IMPROVED DISEASE TRANSMISSION MODEL 33
4.4 SUB-LOCATION MODELING 34
4.5 NEW OUTPUT EVENTS 34
4.6 ENHANCED PRE- AND POST-PROCESSING 35
4.7 COMPUTATIONAL IMPROVEMENTS 38

5 VERIFICATION 39

6 RESULTS: EPIDEMIC SIMULATION WITHOUT DISEASE INTERVENTIONS 42

6.1 EPIDEMIOLOGIC RESULTS 42
6.2 OBSERVATION OF POWER-LAW MIXING 53
6.3 GEOSPATIAL EPIDEMIC DYNAMICS 56
6.4 HOSPITALIZATION BED SURGE CAPACITY ANALYSIS 60

7 RESULTS: EFFECTIVENESS OF DISEASE INTERVENTION STRATEGIES 69

7.1 CASE 1: 2% ANTI-VIRALS, NO PRE-VACCINATION 69
7.2 CASE 3: 4% ANTI-VIRALS, NO PRE-VACCINATION 71
7.3 CASE 2: NO ANTI-VIRALS, 20% UNIFORM PRE-VACCINATION 71
7.4 CASE 4: NO ANTI-VIRALS, 40% UNIFORM PRE-VACCINATION 72
7.5 CASE 5: NO ANTI-VIRALS, 20% TARGETED PRE-VACCINATION 73
7.6 CASE 6: MASKING 75

8 KNOWN ISSUES AND FUTURE WORK 77

9 REFERENCES 81

1 Introduction

This Los Angeles Case Study represents a milestone in the developmental progression of EpiSimS. A set of new epidemic modeling capabilities has been implemented and verified, so that disease dynamics can be simulated when various outbreak control strategies are employed. A detailed model of pandemic influenza has been implemented into EpiSimS, adding to the capability to model smallpox and plague. EpiSimS has been extended to model targeted vaccination strategies, contact tracing with antiviral medication, and transmission reduction by behavior modifications, such as wearing of masks.

A synthetic population has been constructed which includes over 16 million individuals, representing the 2000 US census data for the five counties of Los Angeles, Orange, Riverside, Ventura, and San Bernardino. These counties, plus San Diego county, have been represented physically in EpiSimS as 562,452 *locations.* There is a rough correspondence between EpiSimS locations and US Census blocks (some census blocks map to locations, some map to more than one location).

Several computational advances have been implemented into EpiSimS, resulting in significant gains in efficiency. A new algorithm has reduced the time required to distribute individuals and locations to processors from 4 hours to 45 minutes, for the Los Angeles scenario. By aggregating locations into links, the simulation can more efficiently assign persons to sub-locations (i.e. mixing rooms). Several new scripts rapidly extract select data from the large EpiSimS output event files. Several synchronization strategies have been examined and quantified. EpiSimS has been successfully ported to a 64-bit architecture cluster.

There are two main results sections. The first describes the EpiSimS simulation results for an influenza pandemic in which no effective vaccine or antiviral treatments are available. The simulation results are presented from an epidemiological point of view. A detailed analysis of the hospitalization rates is presented. The second group of results examines the effectiveness of various outbreak response strategies, including contact tracing with antiviral treatment, various targeted strategies of delivering vaccination, and wearing of masks.

2 EpiSimS Simulation Set-up

2.1 Construction of the Synthetic Los Angeles Population

The population is constructed from the 2000 US Census data, counting those people that reside in households. The synthetic population currently does not include the ~3% of the real population that reside in institutions or other group quarters. The included counties and the number of individuals represented in each are shown in Table 2.1-1. The number of census tracts in each county is also given.

County	Population	Census tracts
Los Angeles	9,366,843	2028
Orange	2,812,102	573
San Bernardino	1,672,705	241
Riverside	1,514,716	342
Ventura	740,166	154
Total	16,106,535	3338

Table 2.1-1. Breakout by of the number of individuals in the synthetic population

Prior to running simulation scenarios on Los Angeles, population demographics, people's daily schedules, the initial health state of each person, daily activity participation per location, the disease manifestation and treatments, and an array of scenarios must be assembled in a form usable by EpiSimS. This information is contained in Demographics, Schedule, Health, Partition, Disease Model, Scenario, and Configuration model files. The original Los Angeles population data comes from UPMoST as Person, Household, Activity, and Location entity files and is further processed, filtered, and reformatted for EpiSimS. The model files were created containing the population demographics, people's daily schedules, the initial health state of each person, daily activity participation per location, the disease manifestation and treatments, and an array of scenarios. The creation order and dependencies of each of these model items is shown in Fig. 2-1. The software tools used are also noted.

This collection of model files and a set of run-time configuration parameters comprise the configuration for one simulation run. The Disease Model, Demographics, Schedule, and Partition stay constant over all runs. The initial health states of the population can vary. Different Scenarios are defined for each run based on the disease and interventions being explored.

A Sample facility is available to create a subset of a larger population for testing and debugging purposes, resulting in subsets of the original Demographics, Schedule, Health, and Partition model files. Further description of the individual model elements follows.

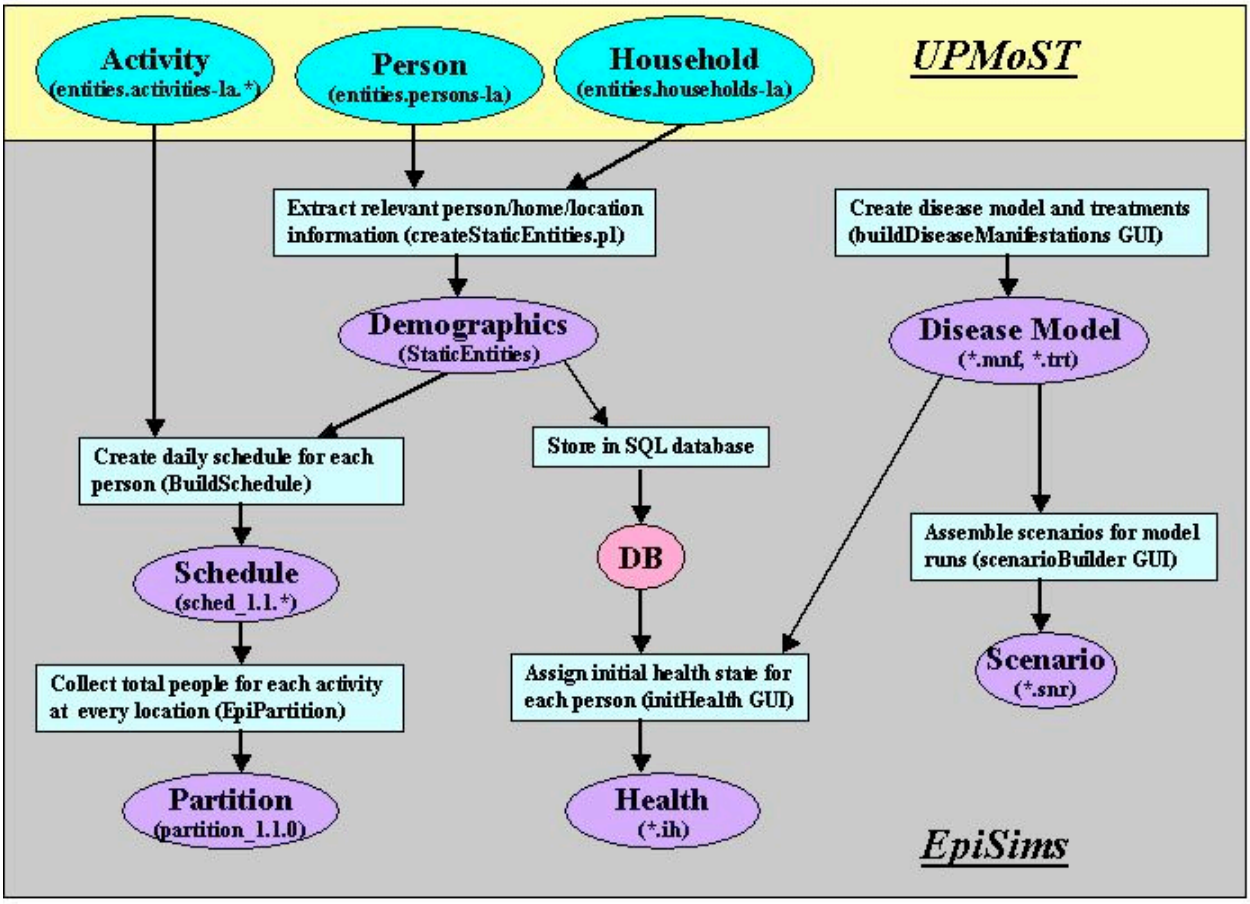


Fig. 2-1. Generation of model files required for run configuration.

UPMoST Population – Person, Household, and Activity Data

The UPMoST team generated the synthetic LA population of over 16 million people, composed of person, household, location, and activity data in the form of entity files from the 2000 US Census Bureau Data, the September 2004 Dun and Bradstreet Business Database , the National Household Travel Survey, the National Center of Educational Statistics, and January 2004 NGA Emergency Response Services. Each entity of a particular type has a unique identification number. The LA synthetic population represents Ventura, Los Angeles, San Bernardino, Orange, Riverside counties, with some work locations in San Diego county.

Person data (entities.persons-la file) contains age, gender, working status and household ID. Household data (entities.households-la file) contains the number of people in a household, number of workers, household income, household location ID, household link ID, state code, county code, tract code, and block group code. A link represents an aggregation of locations. Activity data (entities.activities-la.* files) contains the type, start/stop times, location ID, link ID, and transportation mode for each activity in a person's day. UPMoST activity types include work, school, retail, other, serve passenger, home, college, daycare, visit, social, service, meal, and medical. UPMoST also generates location data with additional zone, state, county, tract, block group, block, and block key information.

EpiSimS does not use the UPMoST entities directly; they are processed into the model files required for running simulations.

Demographics Generation

UPMoST Person and Household information for LA was used to generate demographic data for each person in the simulation. This includes each person's ID, age, gender, worker status (worker or not), household ID, household location ID, household link ID, and household income. The Perl script, createStaticEntities.pl, was used to extract and combine the relevant UPMoST data into a StaticEntities file. This contains all the possible synthetic individuals available for a simulation. It is also used by other generation facilities.

There are 16,106,535 people available for a simulation. An example of part of the StaticEntities file follows. The first line contains the header.

```
# ID TIME TYPE X Y Z HHID AGE GENDER WORKER HHLOC HHLINK HHINC
101 0 P 0 0 0 1 52 1 1 603415 24177694 202000
102 0 P 0 0 0 1 52 2 1 603415 24177694 202000
103 0 P 0 0 0 1 27 2 1 603415 24177694 202000
104 0 P 0 0 0 2 24 1 0 615361 24184499 0
105 0 P 0 0 0 2 22 2 0 615361 24184499 0
106 0 P 0 0 0 2 5 1 0 615361 24184499 0
107 0 P 0 0 0 2 25 1 0 615361 24184499 0
108 0 P 0 0 0 3 48 1 1 604387 24178233 123130
109 0 P 0 0 0 3 44 2 1 604387 24178233 123130
```

Schedule Generation

The LA UPMoST Activity and Demographics information was used to generate a schedule for each person's activities in a day starting at home at midnight till the end of the day. The BuildSchedule program was used to create multiple files containing the schedule information (sched.1.1.* files). A person's schedule contains departure/arrival pairs from one activity to the next continuously through a typical day. These files were indexed by person ID (using Berkeley DB) to allow for fast access to individual schedules. The aggregated locations, or links, were used as the locations for efficiency.

Activities for UPMoST and EpiSimS are not all the same. EpiSimS' activities consist of home, work, school, college, shop, social recreation, visit, serve passenger, and others. The UPMoST activities are mapped into EpiSimS activities (see Table 2.1-2). Some of the UPMoST activities are mapped to the EpiSimS activities because they are similar in characteristics. In EpiSimS, work activities are modeled such that workers always go to the same location and room day after day. The UPMoST daycare activity is mapped into the EpiSimS work activity since children in daycare also typically go to the same location and room. The EpiSimS shop activity is modeled such that a shopper goes to a different shopping area every day. The UPMoST service, meal (dining out), and medical (doctor's appointment) activities are mapped to shop.

UPMoST Activity Type	EpiSimS Activity Type
home	home
work, daycare	work
school	school
college	college
retail, service, meal, medical	shop
social	social recreation
visit	visit
serve passenger	serve passenger
other	other

Table 2.1-2. The UPMoST activity types are mapped to EpiSimS activities for daily schedule generation.

In an EpiSimS simulation, each person moves from location to location participating in different activities throughout a typical day defined by their schedule. People can deviate from their schedules based on the scenario defined. For example they will self-isolate at home while they are incapacitated based on their disease state if self-isolation is defined in the scenario.

An example of part of person 101's schedule follows. Person 101 starts off at his home at location 24177694 at midnight. At 6:30 AM he leaves for college, arriving for a class at 6:45. After that he returns home at 8:15 and then leaves for a shopping trip at 10:15.

```
00:00:00 24177694 101 4 1          Person 101 starts off at their home
06:30:00 24177694 101 1 23915724   Departs for a new location
06:45:00 23915724 101 0 8          Arrives at location for college activity (8)
08:00:00 23915724 101 1 24177694   Departs for new location
08:15:00 24177694 101 0 0          Arrives back home for home activity (0)
10:15:00 24177694 101 1 110158614  Departs for new location
11:00:00 110158614 101 0 2         Arrives at location for shop activity (2)
...
```

Partition Generation

Partition information is generated from the schedule information. All the possible locations are enumerated and the total number of people that participate in each activity at each location is collected from all the schedules. The EpiPartition program was used to generate the LA partition file, partition_1.1.0. In the case of LA, we used the aggregated locations, or links, as locations. The simulation uses the partitioning information during initialization to assign the locations randomly to processors when running and to determine the number and type of buildings and rooms created at each location.

The LA partition file contains 562,525 locations. A few lines from this file follows. Each line represents one location. The total number of people participating in each activity per day at each location is shown.

```
                          Activities Available at Each Location
                  home   work   shop   visit  social other  serve  school college
                                                -rec          -pass
24177694      0   15     0      0      1      0      0      0      0      0
23915724      0   123    42     215    33     170    12     63     0      723
110158614     0   49     11     157    5      0      1      19     0      0
110126576     0   249    1349   5902   102    1      57     731    0      0
```

Health Initialization

An Initial Health disease state is specified for every person. The population entity files are imported to a local database so that the population can be conditionally divided into groups dependent on demographics. The Initial Health program also uses the Disease Model manifestations so that the states can be viewed and assigned (see Fig. 2.1-2). Normally, a large percentage of the population is assigned to an uninfected disease state and a very small percentage is assigned an infected state. The LA Case Study started with a total of 202 index cases who were randomly chosen from each of the demographic groups. This subset was put into the latent incubating disease state in the disease manifestation appropriate to their age demographic.

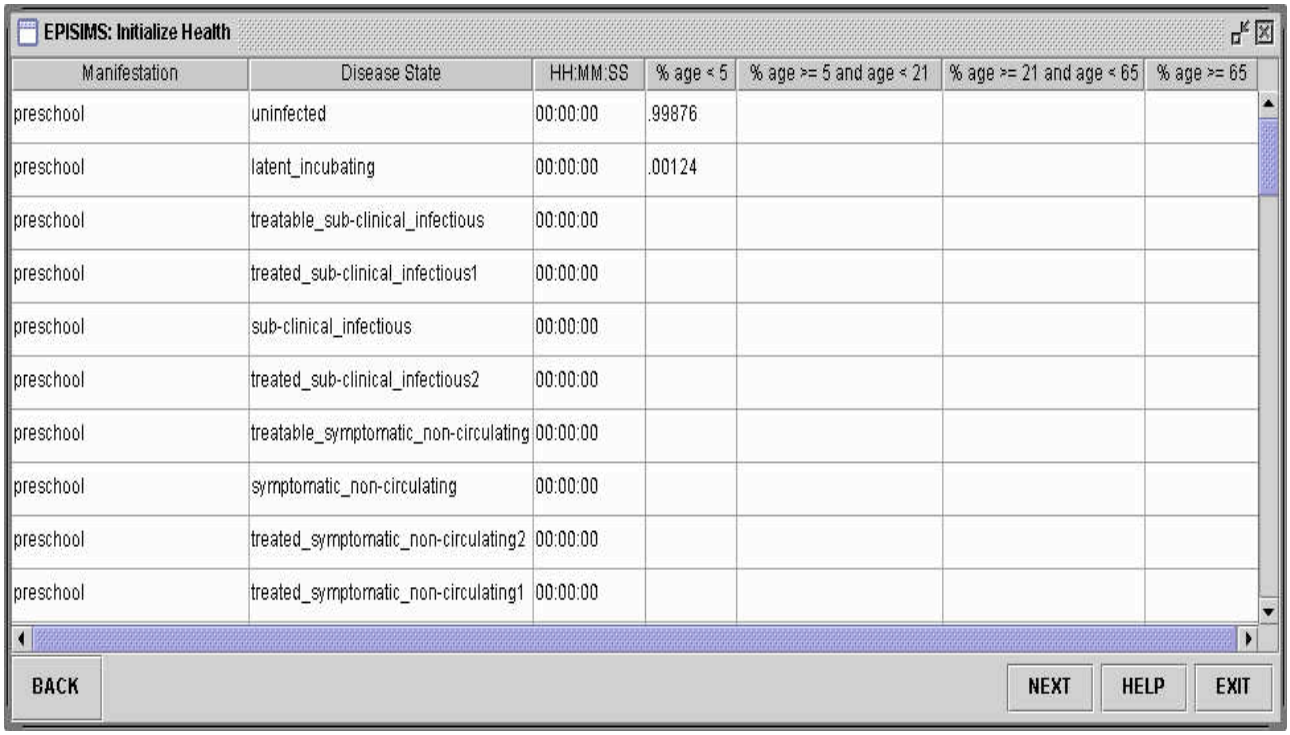


Fig. 2.1-2. The Initialize Health tool was used to assign the LA population's initial disease states.

The time field (HH:MM:SS) indicates the time spent in a disease state prior to the beginning of the simulation. The LA Case Study used times of 00:00:00 in all health files. Several sets of initial health input files were made for the entire population, using different random seeds to get different index cases. In addition, initial health input files were created for all reduced-population sample files.

Sample Generation

Sample generation creates smaller-sized subsets of a city's original population that are a fully connected by the activity locations they share. Samples are generated from the city's original Schedule and Demographics. The set of people in a sample are stored in a database and a subset of the city's schedule is extracted. Samples are typically created of size 1K, 10K, 100K, and 1M to be used for debugging and disease manifestation calibration.

The Sample program is used to extract small populations out of big ones. A randomly selected subpopulation doesn't work well to test EpiSimS, because connectivity decreases as the population becomes smaller and smaller. As the subpopulation becomes smaller, more and more people are walking into empty rooms. Care must be taken to choose people who are connected to others in the subpopulation. This is done iteratively. A location is picked at random. All people who spend time at this location are added to the subpopulation. As people are added to the subpopulation, all others in that person's household are added as well. As people are added to the subpopulation, their schedules are written to a schedule file, and all the locations they visit are added to a heap of locations. When everybody has been added from the targeted location, another location is picked from the heap of locations, and the iteration continues until the subpopulation achieves its desired size.

2.2 Scenario Generation

The EpiSimS configuration file contains the many configuration parameters that are used to set up the scenario and to control the simulation run. The configuration keys are used to specify input and output files and directories, computer architectural details, sublocation modeling, treatment response parameters, mapping of UPMoST activity types, and specific run variables such as number of days to run and how often to capture output. The configuration file is specified along with the executable file when a simulation is invoked.

The configuration file is formatted as key/value pairs. One entry, for example, specifies that EpiSimS set the key *UPMOST_ENTITY_FILE* to point to the file *$data_dir/entities/StaticEntities*, which contains the population demographics data. Additional key/value pairs in the configuration file specify:

- The schedule file that gives the daily schedule of each person in the population.
- The initial health file that gives the initial disease state and resident time in that state for each person and a transmission coefficient group for each person.
- The disease treatments used in this simulation run.
- The disease manifestation file that specifies the disease model.
- The scenario file that lists the exogenous events in the simulation: what treatments will be applied and when; a list of contamination events if any; and when self-isolation, social distancing and masking behaviors begin.
- The partition file that lists the number of people at an activity at a location. It is either calculated as the total number of people per day or the maximum number per hour in a day.
- A directory for writing output files that record the time and simulation messages. There is a log file created for each processor running the simulation. At the end of simulation run each processor calculates statistics for event handling, synchronization, and CPU usage.
- A set of events output files that record all the disease and treatment statistics. There is an event file written by each processor and the events are sorted and merged in post-processing analysis. The three types of events recorded are exposure, treatment, and disease change.
- A number that specifies (in seconds) how often to print output.
- A set of sub-location modeling parameters giving target mean values for the number of persons in mixing groups of various kinds, such as students per classroom. The user can specify the number of mixing groups of various types, or specify that the partition algorithm compute the number of mixing groups at run time.
- A set of consequence mitigation flags and parameters, specifying which treatment strategies are employed, and giving the fraction of contacts named and the fraction of named contacts

found, and the fraction of time that a mask would be worn, by demographic group and activity.

- The names of headers in the entity file that map UPMoST categories to EpiSimS categories.
- A set of simulation control parameters to select the synchronization method, the slack time, the algorithm used to assign locations to processors, the duration of the simulation, the number of CPU's to use.
- A set of parameters used to generate a dendrogram file representing the computed social network. Which demographic group – activity pairs are assigned to the same room each day, and which go to different rooms each day.
- Whether to log information about specific persons or locations throughout the simulation.
- A set of parameters to specify how a simulation will be saved or re-started.

A contamination event is characterized by the time at which the contamination occurs, the infectivity factor of the contaminated room, which location or locations are contaminated, and which building types are contaminated.

When a room is contaminated two things happen. First, the room moves out of its uninfected disease state into an infected disease state. Second, the room takes on the infectivity factor of the contaminate event. The room's net infectivity is the infectivity factor multiplied by the infectivity of the room's current disease state. Once the room is in an infected disease state, it follows its chain of disease states until it returns to the uninfected disease state. If there is a subsequent contamination event, the room's infectivity factor is changed to the infectivity factor of the latest contamination event.

A user can specify that a contamination event will affect a single location or affect every location in the simulation. The user can also specify that a single building type is contaminated or that all building types are contaminated. This allows the full range from a contamination infecting a single building to a contamination infecting every room in town. If it is desired to infect a smaller geographic area, one can find a list of locations in that area and use one contamination event line per location.

Mass delivery of treatment is characterized by the time that the mass delivery begins, the treatment to be delivered, the number of available doses, the number of treaters, the average time it takes to make a delivery, and the range of ages to be treated. Treatment is delivered randomly, household by household, till the number of available doses have been delivered. If there is more than one delivery of the same treatment, the doses are apportioned relative to the fraction of people in each delivery's age group. For example, say there are two deliveries, one for the young, who comprise 7% of the population, and another for the very old, who comprise 3% of the population. If there are 100,000 available doses of an antiviral, the antiviral would be used up when 70,000 kids and 30,000 very old have been treated.

The number of treaters, on the other hand, are specified per delivery. There might be 200 treaters available to treat the young and 100 to treat the old. Each treater delivers treatment to one household. The time per delivery is the average amount of time that will be spent to deliver treatment to the people in a single household. If an age range is specified, only the people in that age range are treated. If no one in a household is in the age range, a treater will not be sent to that household.

The configuration key EPI_HOME_FRAC_FOUND can be used to decrease the number of people who are treated. When a person is found at home, this key value is compared to a random number. If

it is less than the random number, the person is not treated. When set to true, the configuration key PRE_MASS_TREATMENT specifies that all the available doses are distributed before the simulation begins.

Ring delivery of treatment is characterized by the time that the ring delivery begins, the treatment to be delivered, the number of available doses, the number of treaters, the average time it takes to make a delivery, the range of ages to be treated, the attribute that triggers delivery, and the threshold value of the attribute that triggers delivery. As in mass delivery, the available doses are shared among multiple ring deliveries. In fact, if mass delivery and ring delivery are specified for the same treatment, the available doses are shared between ring and mass delivery. The standard CDC policy is to attempt ring delivery early in an epidemic. If the epidemic gets out of hand, mass delivery is instituted to reach more people quickly. The trade off is that more doses are wasted on people who have not yet been exposed.

Treaters are specified per delivery. They are not shared among ring or mass deliveries. In ring delivery, a treater is sent out whenever somebody triggers delivery, for example, by becoming symptomatic. The treater locates and treats a fraction of those individuals who came in contact with the symptomatic person. Disease states have three independent attributes named prodrome, symptoms, and incapacitation. Each of these attributes can be assigned an integer value. Ring delivery is triggered when the specified attribute reaches a threshold value. For example, if the trigger value of prodrome is 3, a person will be contact traced when his prodrome value becomes greater or equal to 3. By using two scenarios it is possible to trigger a person for contact tracing when his prodrome reaches 3 or his incapacitation reaches 1.

The group of people who will be treated varies by the activity of the target person. For the activities of home and work, every person in the same room with the target person is treated. For school and college, every person in the building at the same time as the target person is treated. For the visit activity, everybody who was in the room on the previous day is treated. No one is treated for the shopping, social recreation, or the activity named "other".

There are configuration keys that modify how many people are found in each activity. The keys EPI_HOME_FRAC_FOUND, EPI_WORK_FRAC_FOUND, and so forth decrease the number of people found for treatment.

Self-isolation is characterized by the time at which self-isolation begins, the attribute that triggers it, and the threshold value of the attribute that triggers it. Disease states have three independent attributes named prodrome, symptoms, and incapacitation. Each of these attributes can be assigned an integer value. Self-isolation is triggered when the specified attribute reaches a threshold value. For example, if the trigger value of symptoms is 3, a person will go into self-isolation when his symptoms value becomes greater or equal to 3. When a person is scheduled to move to a new location and he is triggered for self-isolation, he will go home instead of going to his next destination. When his disease state attributes fall below the trigger values, he will once again resume his ordinary schedule.

Mask alert events are characterized by the time at which the mask alert occurs, and the name of the mask which is recommended. The user supplies a list of masks in a mask file. Each mask listed in the mask file has a name, two fractions which specify effectiveness, and the minimum age for of this mask. The effectiveness is measured by the ability of the mask to protect the user from infection by

others and by the ability of the mask to protect others from the user. The effectiveness of various types of masks are listed in Table 2.2-1.

Mask name	EffectivenessToOthers	EffectivenessToSelf	MinAge
bandana	0.1	0.1	6
dust_mask	0.13	0.13	6
dentist	0.34	0.34	6
surgical	0.50	0.50	6
sick_patient	0.50	0.50	6
N95	0.97	0.97	6
N95+	0.0	0.97	6

Table 2.2-1. Mask effectiveness.

The scenario input are the exogenous events that occur for different runs of the study. Throughout all the runs for this case study we specified a "self-isolation" event would occur whenever a person had an incapacitation of 1. This was necessitated by the Disease Model which differentiated between circulating and non-circulating disease states.

The treatments are also triggered by specifying the corresponding scenario input. A mass delivery of vaccine with a specified number of doses and treaters is used to implement a 20% vaccination strategy, for example. Targeted delivery of treatment is specified by including age range parameters for delivery. Ring delivery of antivirals is specified by triggering on a disease state attribute. For this study we used Symptom >= 1 and started all treatments at time 00:00:00.

The scenario development and specification is dependent on the Disease Manifestations and the Simulation model capability. It reflects the design of the Case Study. The scenario files can be created by an editor or using the Scenario Builder GUI tool, shown in Fig. 2.2-1.

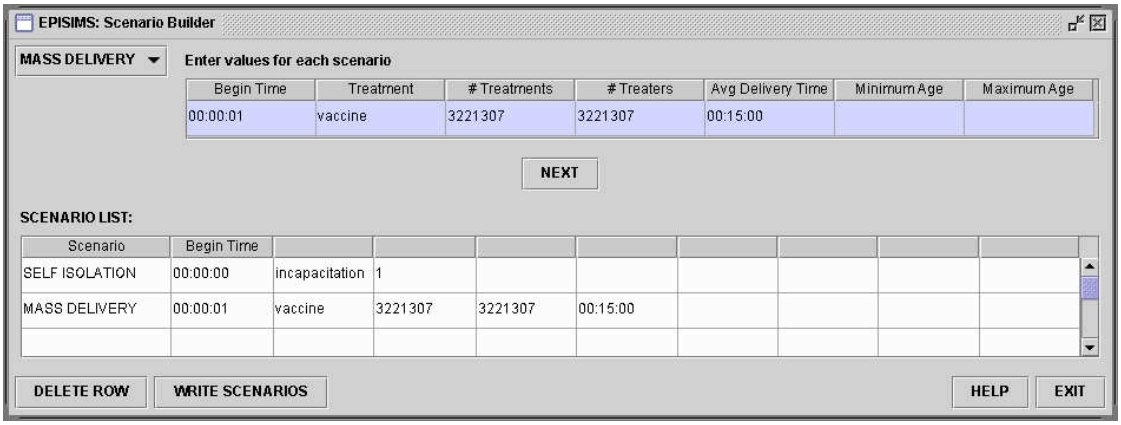


Fig. 2.2-1. The Scenario Builder tool can be used to create scenarios, such as the scenario with 20% uniform vaccination and self-isolation.

3 Pandemic Influenza Disease Modeling

3.1 Influenza Disease Progression

For pandemic type-A influenza, the WHO pandemic planning scenario takes a cumulative clinical attack rate of 25% of the population, in one or two waves of approximately 12 weeks each. The 1918-1919 H1N1 Spanish flu pandemic caused 20-40 million deaths worldwide, with some estimates over 100 million. It infected ~23% of the UK population [UK Dept of Health Plan, 2005]. The 1957 H2N2 Asian flu caused 2M deaths worldwide, and the 1968 H3N2 Hong Kong flu caused 1M deaths. In normal flu seasons, epidemic influenza typically infects 5-10% of the general population, and accounts for 30-40 thousand fatalities in the US, primarily among the elderly.

The following presents a brief overview of information relevant to the modeling of avian-related pandemic influenza in EpiSimS. The basic progression of an influenza infection is through an incubation stage, followed by a symptomatic stage, followed by recovery or death. An infected person is most contagious in the two days immediately following the onset of symptoms. In addition to this basic progression, there are variants in the manifestation of influenza, including a subclinical manifestation and several types of complications.

There are three categories of data to support estimates of the disease characteristics of the next pandemic flu: historical flu pandemics, normal epidemic flu seasons, and human cases of avian flu acquired directly from birds. Historical pandemic influenza typically displayed higher infectiousness, more rapid disease progression, and higher fatality rates among young adults than seasonal epidemic influenza.

Incubation Stage

For influenza types A and B, several sources [Merck Manual, virology-online.com/ viruses/ influenza.htm] state that the incubation stage lasts 48 hours. However, an examination of historical case records [Longini 2004] found that the incubation period of epidemic influenza ranges from one to three days (30% of cases incubate for 1 day, 50% for 2 days, and 20% for 3 days). The average incubation stage duration corresponding to this histogram is 1.9 days.

The EpiSimS stage duration histogram is formulated with intervals all of the same duration. To best accommodate the one-day histogram data, it was most appropriate to formulate a histogram in terms of the fraction of cases that transition during half-day intervals. The incubation stage sojourn time distribution is described by the half-day histogram {0, 0.12, 0.18, 0.259, 0.238, 0.13, 0.07, 0.003}, giving respectively the fraction of cases that incubate for a period of between 0 and 0.5 days, 0.5 and 1.0 days, etc. before transitioning to the infectious stage.

The half-day histogram gives an averaged incubation stage duration of 1.9 days, which matches the one-day histogram exactly. The validity of the half-day histogram is further checked by comparing the resulting cumulative distribution function with that of the integral-day histogram. This comparison is shown in Fig. 3.1-1, where it can be clearly seen that the EpiSimS half-day histogram is consistent with the original histogram.

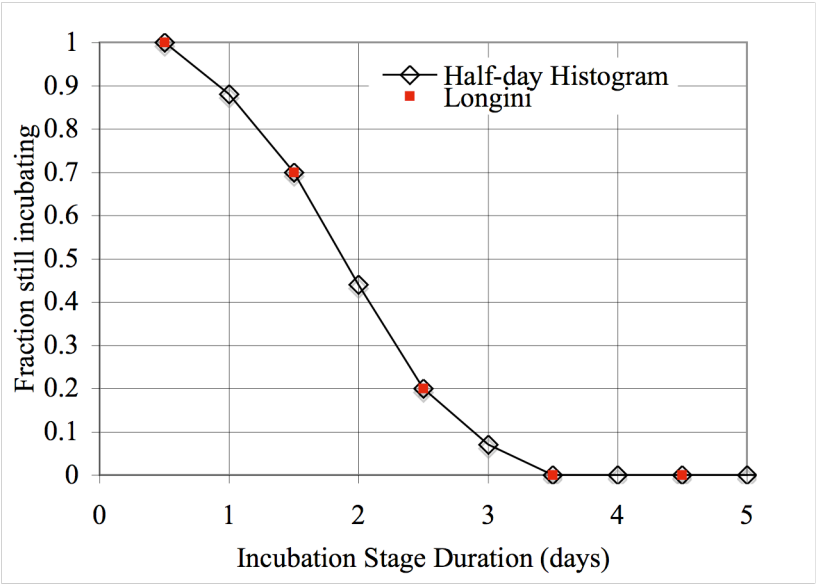


Fig. 3.1-1. Fraction of cases still incubating/latent, as a function of time after infection. The solid squares represent the integer-centered, one-day histogram values from Longini. The diamond markers show the CDF for the half-day EpiSimS formulation histogram.

This half-day histogram replaces a previous EpiSimS implementation of influenza in which the incubation/latent stage duration was uniformly distributed from 1.0 to 3.0 days.

Symptomatic Stage

The first symptoms are headache, chills and a dry cough, followed by fever. The 2005 UK influenza plan asserts that the symptomatic stage duration is four to five days for adults, two weeks for children, and three weeks for immunocompromized persons. Anecdotal reports claim that even though symptoms may last 2 or 3 weeks, cases are only infectious for the first few days of symptoms. Based on examination of historical case data, [Longini 2004] gives an average infectious stage duration of 4.1 days, with 30%, 40%, 20% and 10% of symptomatic cases having symptomatic stage durations of 3, 4, 5 and 6 days, respectively. [Hyman, 2003] also gives an average infectious period for H3N2 of 4.1 days.

Longini's data has been used to construct a half-day resolution histogram for use in EpiSimS, as follows: The infectious stage sojourn time distribution is described by the histogram {0, 0, 0, 0, 0.005, 0.125, 0.16, 0.205, 0.205, 0.12, 0.08, 0.06, 0.04}, giving the fraction of cases that are infectious for 0 to 0.5 days, 0.5 to 1.0 days, etc. The average case is infectious for τ_I=4.1 days. Fig. 3.1-2 shows a comparison of the cumulative density function of Longini's histogram and the half-day EpiSimS histogram.

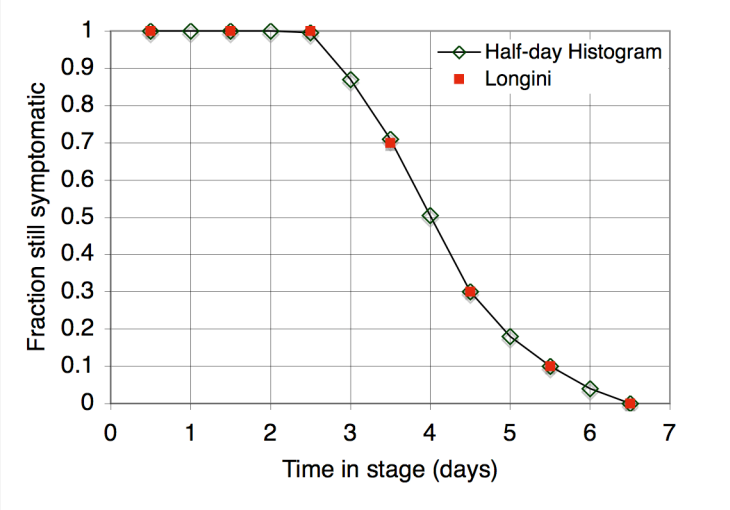


Fig. 3.1-2. The fraction of cases that remain infectious at given times after onset of infectiousness.

A previous EpiSimS implementation of influenza had a symptomatic/infectious stage duration uniformly distributed from 3.0 to 6.0 days.

Subclinical Manifestation

[Elder, 1996] studied 518 unvaccinated health care workers during the mild 1993-94 influenza epidemic in Glasgow. 23.2% were found to have serum antibodies to the influenza strain that circulated during the season, indicating they had contracted influenza. This gives a direct measurement of the attack rate among unvaccinated healthcare workers. Of those that had been infected with influenza: 41% recalled having influenza, 32% thought they had a non-flu respiratory illness, and 27% thought they had no illness. Thus, this study indicates that 27% of infections were sub-clinical in the subject population.

Citing Elder, Longini models that 1/3 of infections are subclinical, and 2/3 are symptomatic. He also asserts that subclinical cases are half as infectious as symptomatic cases. The 2005 UK influenza plan asserts 50% of cases produce no symptoms, and that children are more likely to have subclinical cases. During the "Hong Kong" pandemic of 1968, subclinical infections (producing no symptoms, or symptoms of a mild cold) accounted for 75% of H3N2 cases.

Based on [Longini 2004], the disease model in EpiSimS takes one-third of infections to be sub-clinical. Sub-clinical cases do not exhibit symptoms, but they do become infectious. Sub-clinical cases are taken to be half as infectious as cases that exhibit symptoms. Their incubation and infectious stage durations follow the same histograms as those cases that do exhibit symptoms. Individuals with sub-clinical manifestations continue their normal activities during their "illness".

Complications

There are three main complications associated with influenza. These are 1) bronchitis, 2) primary viral pneumonia, and 3) secondary bacterial or viral pneumonia. About 20% of influenza cases are complicated by bronchitis (or tracheobronchitis or bronchiolitis). Diabetics, persons with ischemic heart disease, and persons over 60 are especially susceptible to bronchitis-related complications.

Pneumonia refers to many diseases that involve infection or inflammation of the lungs. [American Lung Association Pneumonia fact sheet]. In normal epidemic flu seasons, most flu-attributed deaths are caused by secondary pneumonia. The 1918-1919 pandemic influenza, however, also killed its victims through primary viral pneumonia (i.e. pneumonia caused directly by the influenza viral infection) [Barry 2004].

Secondary bacterial pneumonia co-infections (caused by several species of bacteria, including Staphylococcus aureus, S. pneumoniae, and Haemophilus influenzae) usually occurs late in the course of the disease when the primary viral infection is abating [virology-online.com/ viruses/ influenza.htm]. Flu vaccination cuts risk of pneumonia complications in half. In the 1957 pandemic, 28% of cases with staphylococcal pneumonia died, while only 12% of non-staphylococcal pneumonia cases died [UK Health Departments' influenza pandemic contingency plan annex C.] Deterioration can be so rapid that persons entering the hospital die within 48 hours, so antibiotics did not get administered in time. In the 1957 pandemic, the fatality rate of cases with pneumonia dropped from 20% to 13% as doctors realized that pneumonia had to be treated immediately.

About 50% of influenza-associated secondary pneumonia cases are viral. There are no treatments for viral pneumonias. There is a viral-pneumococcal vaccine, which is 80% effective in adults, but less effective in high-risk demographics.

Transmission

Transmission is person-to-person via inhaled aerial droplets and fomites (i.e. moist material in bedding or other cloth). 10% of cases begin shedding virus "just before the onset of symptoms" [UK pandemic plan], while the remainder begin shedding after onset of symptoms. Alternatively, the CDC and WHO influenza fact sheets give that infected persons may be contagious for up to one day prior to onset of symptoms, although at a lower level of contagiousness.

The infectiousness during the symptomatic/infectious stage is age dependent. The baseline infectiousness (i.e. the probability per hour that a particular susceptible person will become infected, given that there is one symptomatic infectious adult or senior in the same room at the same time) is 0.00285 transmissions per hour. This baseline infectiousness was selected to give an epidemic that infects about 25% of the population. The probability that a susceptible person becomes infected during a visit to a room depends on: how many infectious persons co-occupy the room, how long each contact lasts, the type of activity, and the infectiousness category of the infectious person [Eubank 2004]. Symptomatic children and preschoolers have twice the baseline infectiousness, i.e. 0.0057 transmissions per contact-hour. Subclinical persons have half the infectiousness of symptomatic persons in the same demographic.

Since EpiSimS takes infectiousness input as transmission probability per infectious person per symptomatic person per minute, the EpiSimS infectiousness input values are 0.0000475 per minute for symptomatic adults and seniors, 0.00002375 for subclinical adults and seniors, and 0.000095 for symptomatic children and preschoolers.

Self-isolation and Hospitalization

The [Elder 1996] study found that only 48% of healthcare workers infected with influenza took sick leave. [Longini] asserts that 80% of preschoolers withdraw to the home when symptomatic, as do 75% of school agers, and 50% of adults. Persons who withdraw to the home then can only infect

household members. EpiSimS models that half of symptomatic adults and seniors, 75% of symptomatic school-agers, and 80% of symptomatic pre-schoolers will remain home during their symptomatic stage. These "self-isolators" would thus not transmit the disease to anyone except members of their own household and visitors to the house. Those symptomatic persons that do not remain at home are designated *circulating*, and continue with their normal pattern of contacts.

[Halm 2002] gives ~4M cases of community-acquired pneumonia per year in US, leading to 1M hospitalizations per year, costing 9B$/year (average 9k$/admission). On average, there are 11.375 restricted-activity days per case, and 6.56 bed-days per case.

A baseline flu-season activity level is defined as 0.03% of population seeking flu-related new general practician consultations per week [UK influenza plan, 2005]. During a pandemic, this grows to at least 0.5% per week. The 1969/70 pandemic hit a peak of 1.26% new consultations per week for flu.

The decision to admit a patient to a hospital is based on the Pneumonia Severity Index, which assesses age, disease history, and vital signs to rank cases into low (risk levels 1, 2 and 3: 0.1, .6 and 0.9% mortality), moderate (risk level 4: 9.3% mortality), and high risk (risk level 5: 27% mortality). The admission decision is not based on whether the pneumonia is bacterial or not. All moderate and high risk cases are admitted. 43-58% of low risk patients are admitted. There is an option for 23 hour inpatient observation of risk level 3 patients, during which they receive antibiotics and hydration. The median time to clinical stability and discharge is 3 days for low-risk patients, 4 days for moderate-risk patients, and 6 days for high-risk patients. Clinical stability is defined such that when a patient recovers to this level, there is a 1% risk of serious clinical deterioration. Clinical stability means that the fever is resolved, respiratory symptoms are improving, and the patient can take oral antibiotics. Patients still feel sick and may require weeks to return to normal activity after discharge from the hospital.

There are 19k to 193k (mean 95k) hospitalizations per year in the US for which the primary hospital discharge category is listed as pneumonia or influenza, and which are associated with influenza virus infections. An additional 39k hospitalizations per year were influenza related, but had P&I listed as a non-primary discharge category. The total is 134k influenza-related hospitalizations per year (< 0.5 per 1000 population per year). Seasons dominated by influenza A(H3N2) give more hospitalizations than other strains. Persons aged 85+ were hospitalized at a rate of 11.95 per thousand hospitalizations per year for P&I. Children under 5, and older adults (50-64) had 1.08 per thousand, hospitalizations per year for P&I. [Thompson].

Fatality

The mortality worldwide for the Spanish flu (1918-1919) was 3 per 1000 of population in 1918, and a further 1.17 per 1000 in 1919. The 1957 flu had mortality of 2.3 deaths per 1000 population. 2/3 were in persons over 55. [UK Health Departments' influenza pandemic contingency plan annex C.]. For "normal" epidemic influenza, the age-averaged overall case fatality rate is 0.37% [UK influenza plan. 2005] (at 25% attack rate, a pandemic with .37% case fatality would have a mortality rate of 0.925 deaths per 1000 of population.)

In the outbreak in 1997 in Hong Kong, 18 persons were infected with influenza A(H5N1) through contact with birds. 6 died and 3 were severely ill. The H5N1 strain (avian) caused 97 confirmed

human cases in the 2003-2004 and 2004-2005 seasons. 53 were fatal: case fatality rate was 54.6%. All of these cases were probably contracted directly from birds.

In cases complicated by pneumonia, the 30-day mortality rate (case fatality rate) is 5.1% in ambulatory out-patients and hospitalized patients, and 36.5% in patients requiring intensive care.

For seasonal influenza, Longini gives the case fatality rate as 2% for seniors, 0.0294% for adults, 0.0021% for school-agers, and 0.00263% for preschoolers. In order to simulate an avian influenza pandemic, based on the 1918 Spanish influenza, the EpiSimS pandemic influenza model specifies a case fatality rate of 2%, independent of age.

EpiSimS Untreated Disease Manifestations

The basic structure of the influenza disease manifestation consists of 7 states: uninfected, latent/infected, subclinical/infectious, symptomatic/circulating, symptomatic/not-circulating, dead and recovered. The flow paths between nodes is shown in Fig. 3.1-3 This configuration is the same as that given in Longini, and that was used in the previous EpiSimS influenza implementation.

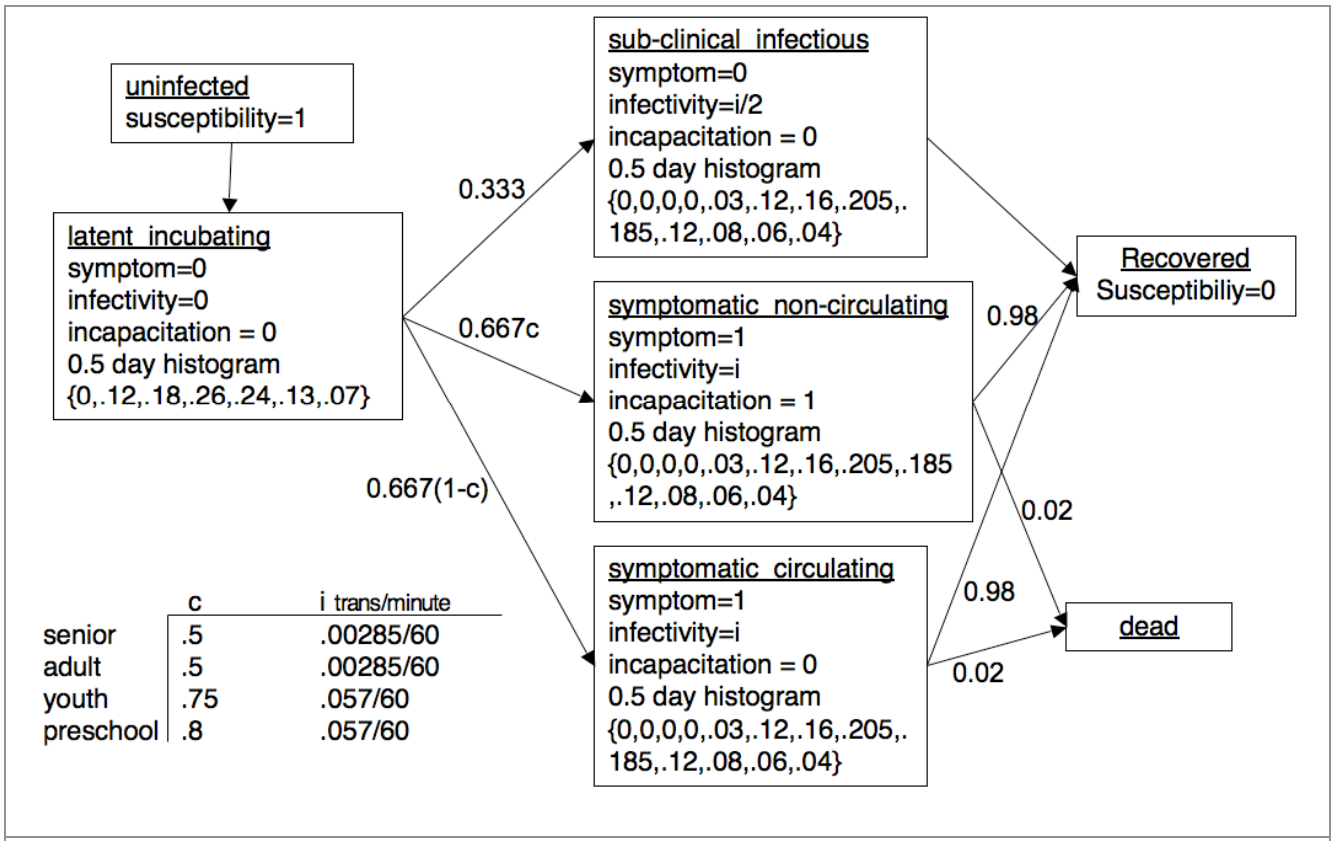


Fig. 3.1-3. Epidemic influenza disease manifestations for four age demographics, with no treatment. The manifestations for different demographic groups is obtained by using the appropriate self-isolating fraction and infectiousness values.

A separate disease manifestation of the form shown in Fig. 3.1-3 is implemented for each of four demographic categories: Pre-school (age < 5), Youth (age >= 5 and age < 21), Adult (age >= 21 and age < 65), Senior (age >= 65).

The fraction of symptomatic persons that stay at home is: 50% of adults and seniors, 75% of Youth, and 80% of Preschoolers. The base infectivity levels in the disease manifestations are: 0.0000475 transmissions per minute per contact for symptomatic adults and seniors, 0.000095 transmissions per minute per contact for symptomatic youth and pre-schoolers. For the sub-clinical infectious state, the infectiousness is half the base infectivity level.

3.2 Vaccination and Antiviral Medication

The Efficacy of Influenza Vaccine

For seasonal influenza, influenza vaccine is typically 70% effective in preventing infection [Longini 2004]. This number combines 1) the effectiveness of the vaccine against the three strains in the vaccine, and 2) the presence of strains that are not included in the vaccine. Those that are vaccinated and do become infected have their infectiousness reduced by a factor of 5, relative to unvaccinated cases (the infectiousness efficiency of vaccination is 79-80%). In addition, vaccination reduces the infectious stage duration by one day.

The Efficacy of Antiviral Medications

Over-the-counter medicines are widely used to treat symptoms. Antibiotics do not affect the virus, but may prevent or cure bacterial infection, especially bacterial pneumonia. There are four antivirals approved by FDA for influenza:

- amantadine hydrochloride (Symmetrel),
- rimantadine (Flumadine), also an amantadine
- zanamivir (Relenza), a neuraminidase inhibitor
- oselatamivir phosphate (Tamiflu), also a neuraminidase inhibitor

Symmetrel and Flumadine are approved for treatment and prevention of influenza A, although CDC recommends its use only for prevention to avoid viral resistance. Symmetrel and Flumadine are ion channel blockers that target the M2 viral protein. However, the H5N1 strain (avian flu) has two mutations in the M2-producing gene making it resistant to adamantanes. Amantadine may prevent influenza if taken continuously at the time of an epidemic. It is not widely used against epidemic influenza, and is usually restricted to high-risk persons. Antivirals, especially adamantanes, lead to resistant strains of influenza [FDA] when used as preventative, and have adverse side effects.

Tamiflu is approved for treatment and prevention of influenza A and B. Tamiflu is available in pill form. Tamiflu is more effective against some of the 9 NA types than others. It is most effective against N2. To achieve the same effectiveness, 10-30 times as much dose is required for N1 than for N2 types.

Relenza is approved for treatment of influenza A and B. Relenza comes in the form of a nasal spray, and is not widely available.

Treatment with antivirals is recommended to begin within 48 hours of the onset of symptoms. All four antivirals typically reduce the symptomatic period by 1-2 days if given within 48 hours of symptoms. The British plan is to stockpile enough Tamiflu to keep vital services and health care functional. Similarly, the CDC recommendations give highest priority for antivirals to workers in nursing homes, hospitals, and facilities caring for the immunodepressed.

Stockpile and Delivery of Antiviral Medication

The Strategic National Stockpile, operated by CDC has large quantities of medicine and supplies for emergencies severe enough to cause local supplies to run out. Upon commitment, supplies can be delivered to any state within 12 hours. Each state is responsible to receive and distribute supplies to local communities. Their objective is to be able to supply several large cities simultaneously, particularly in the event of terrorist attacks or large-scale natural disaster. Each state or local community is responsible to provide information on how to get medicine via TV, radio, newspaper, internet, etc.

Some of the supplies are organized into *Push Packages*, which are pre-positioned in strategically located, secure warehouses. In addition, the SNS uses vendor managed inventory (VMI) for follow-on supplies that might be needed within 36 hours. Local stocks of medical materiel will be used for first response, then SNS will bolster the supplies with a combination of 12-hour push packages and VMI. Push Packages are accompanied by technical advisory response unit staff.

The Strategic Reserves/Stockpiles was initiated in 2004 with 80M$ of funding. DHS invested 40M$ in 2004 and 40M$ in 2005 to stockpile 4.5M doses of influenza vaccine in the Vaccines for Children Program; 87.1 M$ to stockpile 2.3M regimens of tamiflu; and 34M$ on Rimantadine capsules to treat 4.25M adults and Rimantadine syrup to treat 750,000 children. The production rate of Tamiflu (made by Roche, Hoffmann) has been ~1.5M doses/courses per year. Roche plans to increase production to 4M courses per year, where a course is 10 pills.

For the 2004-2005 flu season, Aventis Pasteur produced ~50M doses of flu vaccine, where ~100M doses were anticipated to be needed. After the shortfall was identified, the distribution of the remaining 22M doses were prioritized to:

- state and local health departments
- the Vaccines for Children Program
- children's providers
- dept of veterans affairs and Indian health service
- long-term care facilities and acute care hospitals
- the Visiting Nurses Association of America
- DOD.

The CDC guidelines for use of antivirals provides for their use during normal flu seasons to protect persons at risk for complications. It is not intended to address pandemic flu seasons. In November 2004, The Strategic National Stockpile held ~1.3M regimens of rimantadine tablets, 60,000 regimens of rimantadine syrup, 859,993 regimens of oseltamavir capsules, and 110,336 regimens of oseltamivir suspension. Among antivirals, only oseltamavir/Tamiflu is effective against avian flu, so it is being held in reserve.

In addition to flu vaccine and flu antivirals, pneumococcal vaccine can prevent a common complication of influenza. For the 2004-2005 flu season, Merck tripled its production of Pneumovax.

Treatment with antivirals has four effects that are modeled in EpiSimS: prevention of infection, prevention of symptoms in infected persons, reduction of infectiousness of infected persons, and reduction of duration of symptomatic/infectious stage. If antivirals are given within one day of onset

of symptoms, the duration of the symptomatic stage is reduced by one day. Presumably, the duration of the infectious stage of the sub-clinical manifestation is similarly reduced by one day.

[Longini 2004] models that a susceptible person taking antivirals has 70% as much likelihood of becoming infected as one not taking antivirals. If a person on antivirals does become infected, there is a 60% probability that the disease will not advance to the infectious stage. Persons on antivirals that do become infectious will transmit disease at one fifth the rate of those not on antivirals. These values have been used in the EpiSimS influenza disease model.

Antiviral medication and vaccination treatments have been incorporated into EpiSimS by extending the disease manifestations to explicitly include treated and untreated variants of the disease states shown above for untreated influenza. Thus, each of the three untreated disease states are now accompanied by a similar state for persons having received vaccination, another for persons having received antiviral treatment, and a fourth state for persons having received both vaccine and antiviral treatment. The seven disease states used to implement untreated influenza are thus extended to 16 disease states to allow for antiviral medication and vaccination. This rather awkward solution is needed because the antiviral and vaccine treatments change the stage transition histograms. Each of four demographic groups has a demographic-specific set of disease parameters, but the same structure of 16 disease stages. A diagram of the treatment-extended disease stages is shown in Fig. 3.2-1.

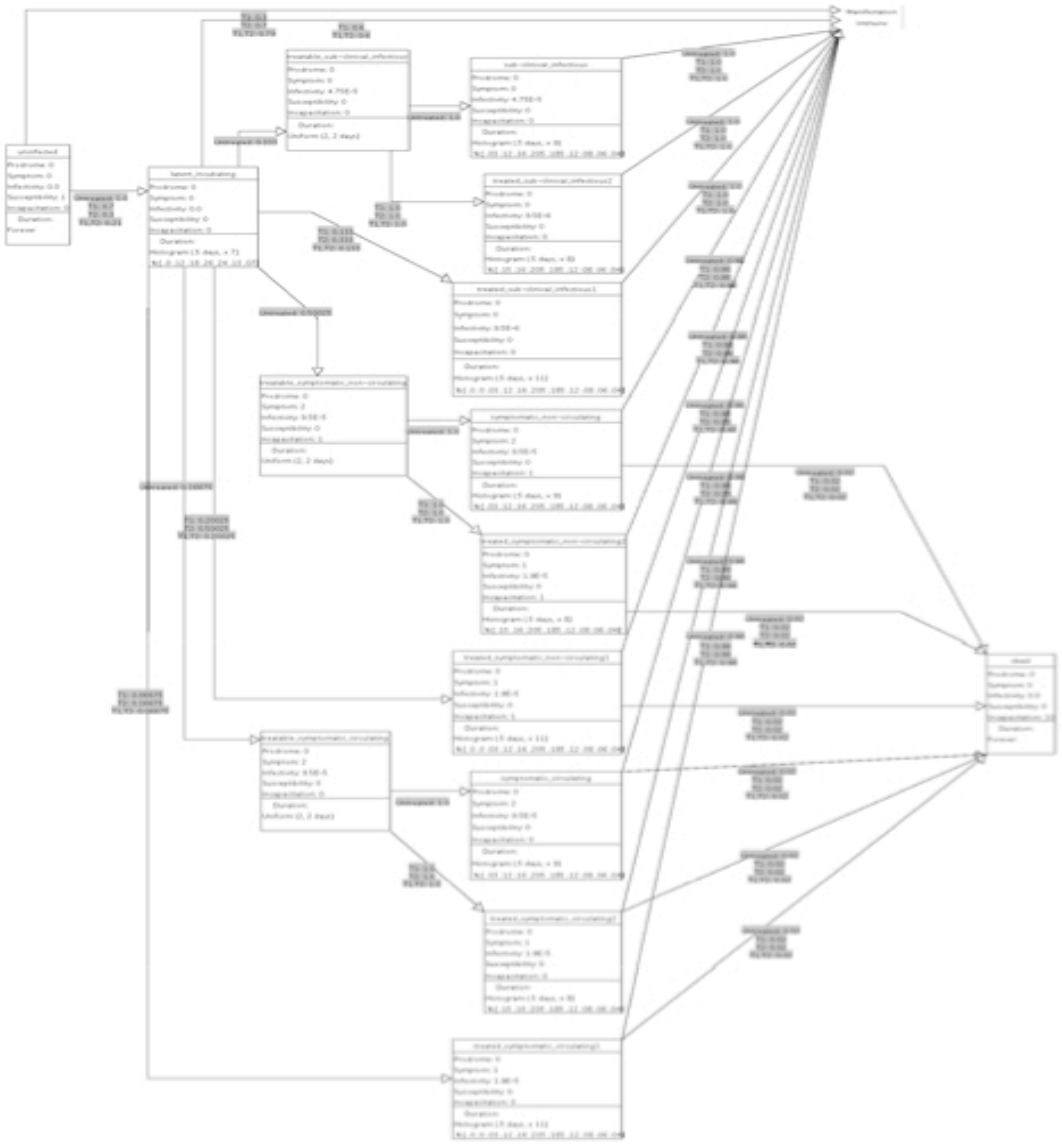


Fig. 3.2-1. Expanded disease manifestation to account for antiviral and vaccine treatments.

EpiSimS reads the detailed characterization of the disease model from a disease manifestation file. A GUI tool, BuildDiseaseManifestation, allows a user to graphically model the disease stages, transition histograms, effectiveness of treatments, and demographic dependences. The tool then automatically creates the disease manifestation file. The BuildDiseaseManifestation tool has an "Adjust Infectivity" function so that infectivity levels can be changed without reconstructing the

entire Disease Model. A partial screenshot of the BuildDiseaseManifestation tool is shown in Fig. 3.2.-2.

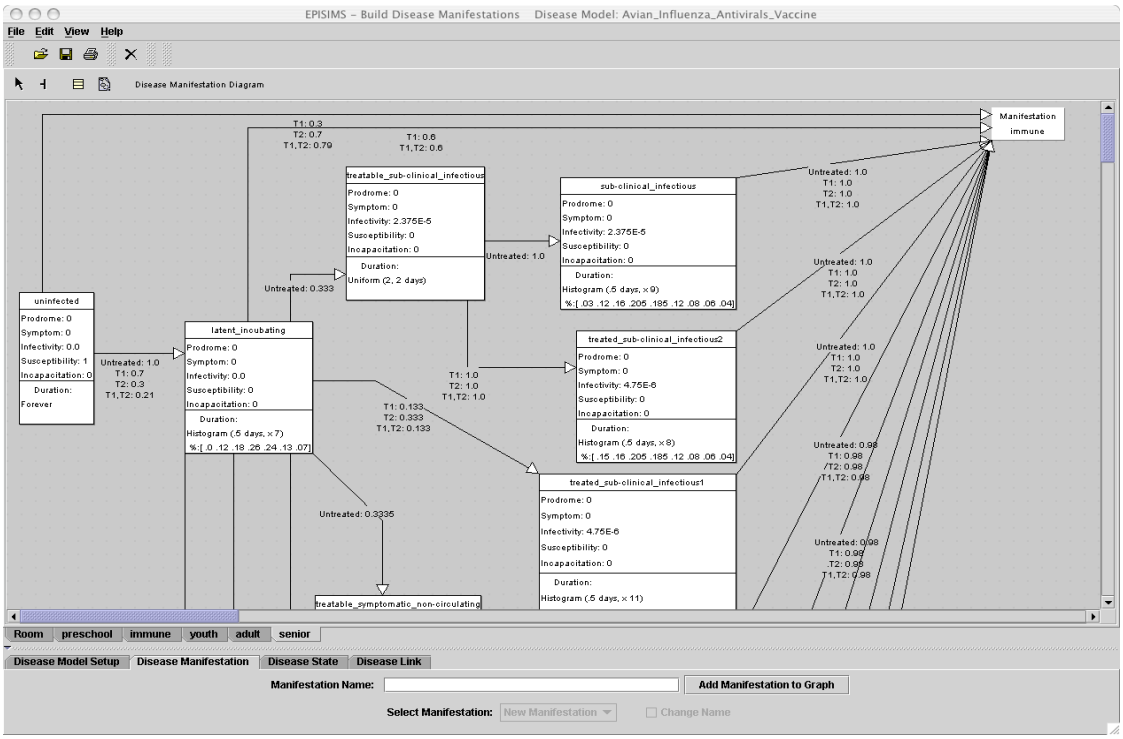


Fig. 3.2-2. The GUI-based BuildDiseaseManifestation tool.

3.3 EpiSimS Infectiousness Calibration

The first milestone in the course of running our disease model on a number of scenarios was to calibrate the infectiousness in the disease transmission model, in order to obtain a roughly 25% attack rate. In particular, the initial values for infectiousness, as first extracted from published literature, were tied to a basic adult-to-adult infectiousness of $i_0 = 0.005$ transmissions per hour per infected person per susceptible person. This basic infectiousness rate applies for household contacts from symptomatic adults or seniors to susceptible adults or seniors. For symptomatic children, the household infectiousness is $2i_0$. For subclinical adults and seniors, the infectiousness is $0.5i_0$, and for household transmissions from subclinical children, the infectiousness is i_0. The transmission occurring in non-household contacts are scaled from these household values. The calibration process determined the appropriate value for i_0 so that EpiSimS obtained a scenario-specified attack rate.

Because it is computationally expensive to run an EpiSimS simulation to completion for high attack rates with more than 16 million simulated individuals, a scoping model (EpiHist) was constructed so that the EpiSimS disease model could be calibrated with short runs (simulating a few weeks). The scoping model implements a non-stochastic uniform-mixing model, and uses the EpiSimS disease progression model with identical stage transition histograms. Scenario parameters, disease progression parameters, and disease transmission parameters can readily be varied. The number of remaining susceptible persons, the number of new infections, the number of persons becoming infectious, and the number of people dying or recovering is computed for each timestep. The number of people transitioning from incubating stage to infectious stage is computed by convolving the new

infections with the incubation-to-infectious histogram of sojourn times. Likewise, the number of people transitioning from infectious to recovered/dead is computed as a convolution of new infectious with the infectious stage sojourn time distribution. The scoping model is implemented in a an Excel spreadsheet which is about 4 megabytes in size and takes about 30 seconds to recalculate, when using timesteps of one hour.

The scoping model takes that *on average* each infectious person would transmit the disease to R_0 new cases, if the entire original population were susceptible. As the fraction of the population that is susceptible decreases, the transmissions per infectious person decreases correspondingly. The transmission coefficient, giving the *average* number of new cases per day per infectious person, is formulated as $\alpha = (R_0 / \tau_I)(S / P_0)$. The first factor R_0/τ_I is the number of transmissions per day per infectious person, where R_0 is the average number of transmissions per case that would occur if the entire initial population of P_0 was susceptible, and τ_I is the average infectious period duration. The second factor reduces this transmission rate to account for a reduction in the fraction of the population that is susceptible.

To illustrate the process of calibrating the EpiSimS infectiousness with short simulations, Fig. 3.3-1 shows the number of new infections per day from an EpiSimS simulation covering only 5.6 days, for a baseline infectiousness reduced to i_0=0.00336 transmissions per hour per susceptible person per infectious person. The scenario shown is for 202 persons infected at time 0, with an initial population of 16.1 million persons. The number of new infections per day computed by EpiHist is also shown, where R_0 has been adjusted to a value of 1.7 in order to best fit the EpiSimS result. We thus conclude that i0=0.00336 per hour corresponds to R0=1.7, when applied to the emergent social contact structure generated by EpiSimS. The scoping model can then immediately run the epidemic to completion, obtaining an attack rate of 70% of the original population.

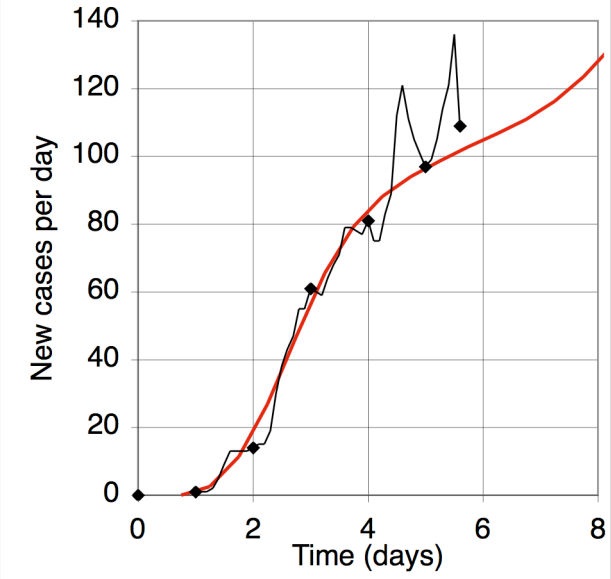


Fig. 3.3-1. Scoping model sample output for the first two weeks of an outbreak, demonstrating fidelity of dynamics on a sub-day timescale that can not be achieved with SIR-type models.

In order to obtain a target attack rate of 25%, the scoping model shows that the value of R_0 should be 1.15. The EpiSimS nominal infectiousness parameter was adjusted so that the few-day EpiSimS simulation gave a new infections per day curve that agreed with the EpiHist curve with R_0 set to 1.15. This value was then used in an EpiSimS simulation of the full epidemic. The resulting simulated epidemic had only a 12% attack rate, rather than the desired 25%.

Close examination of the EpiSimS output discovered that the number of new infections per day per infectious person scaled with the susceptible fraction of the population raised to a power higher than 1.0. This phenomena was explored is some detail, resulting in the formulation of a power-law mixing model to replace the traditional homogeneous mixing assumption that is found in most existing epidemiological modeling. This result is being published in Mathematical Biosciences. For the Los Angeles social contact structure that emerges from the activities of individuals in Los Angeles, the scaling exponent was found to be 2.06. The number of new infections per day per infectious person in the power-law mixing model is then given by $\alpha=q/I=(R_0/\tau_I)(S/P)^{2.06}$. With this formulation, we find that a reproductive number of 1.34 corresponds to an attack rate of 25%. This reproductive number was then found to correspond to an EpiSimS nominal infectiousness value of 0.00285 transmission probability per hour, to give a good match between short EpiSimS epidemic curve and the EpiHist result, and further, that the full epidemic simulation with EpiSimS then gives a 25% attack rate.

4 EpiSimS New Capabilities

4.1 Disease Progression Model Enhancements

Reformulation of the disease progression model implementation

A new disease model was fully implemented this year. The former load-model design had the following drawbacks:

- thresholds affect both infectiousness and progress of the disease, with complicated consequences
- large variability in state duration arising from small variability in initial load
- poor decomposition of infection rate by demographics
- unwarranted precision in model
- profusion of exponentiation and logarithms, slowing computation
- approximates nonlinear dynamics with linear dynamics
- overly deterministic transmission process

The new disease model meets several requirements:

- allow multiple manifestations of disease, possibly based on demographics
- allow contamination of locations
- allow variable infectiousness and susceptibility
- represent effects of a variety of treatments
- relate transmission rate to duration/type of contact and demographics of susceptible and infectious
- allow flexible ordering of symptomatic, infectious, etc.

- allow flexible transmission scaling behavior as number of people increases
- label states with meaningful names (infected, symptomatic, etc.)
- provide easy tracking of state changes
- scale to populations of 10 million
- easy matching between people and the course of their disease
- match state update process to discrete event system
- entire model and scenario specifiable by naive user
- lightweight computation and memory requirements

Different individuals manifest infection by the same infectious agent in different ways. A primary goal of EpiSimS is to capture the dependence of disease manifestation on demographics. To this end, each person in the population is assigned:

- a Disease Manifestation,
- a Disease State and an associated time stamp,
- an (integer) Treatment Level,
- and a Disease Transmission Type.

The assignment is consistent with user-specified probabilities conditional on demographics. A variety of assignments may be produced by varying the random seed. During the course of the simulation, each location (at the finest resolution simulated) is also associated with a disease state. Some aspects of a location's disease state may be modified by exogenous events created by the user (e.g. contamination, decontamination); others reflect transmission dynamics internal to the simulation.

In addition to the conditional probabilities for the assignments described above, the user must also specify the following:

- a set of Disease Manifestations,
- a set of Transmission Rate functions,
- a set of (integer) behavioral threshold values.

Each possible Disease Manifestation must be specified by the user prior to using EpiSimS. A Disease Manifestation is a Markov Chain consisting of a finite set of Disease States together with transition probabilities among them and a distribution of residence times in each state. Most Disease States are associated with values for the following attributes relevant to the spread of disease:

- prodrome, (integer) (non-specific symptoms, easily mis-diagnosed)
- symptoms, (integer) (depending on thresholds)
- infectivity, (floating point) (capable of transmitting disease - may also represent contamination)
- incapacitation, (integer) (cannot perform some activity - distinct from symptoms)

These attributes take on nonnegative values representing the degree to which the attribute is present. The default special value *zero* is interpreted as a complete absence of the attribute; the special value *dead* can be assigned to the incapacitation attribute. The value of some attributes affects the

dynamics of disease transmission directly. For example, infectivity is related to the probability of transmission as explained below. Others may affect the behavior of an infected person: symptom levels reflect the severity of symptoms, the likelihood of health-care-seeking behavior, and the likelihood of correct diagnosis; the degree of incapacitation affects whether a person stays home from work, shopping, or other activities. Detailed interpretation of these attributes is provided below. In addition to these attributes, the user can assign each state a unique alphanumeric name. Simulation outputs and analysis tools will refer to states by this name.

Each Disease Manifestation's Markov Chain must end with one of two special kinds of Disease States. Any Disease State with "incapacitation" set to *dead* is a possible terminal state. The other special state is the *uninfected* Disease State, with the single attribute "susceptibility" used to specify both how likely an uninfected person is to become infected and whether a person who has recovered can be re-infected. Assignment of any state besides *uninfected* implies that the person is infected.

Note that the Disease State does not contain a "recovered" attribute. The simulation maintains information about each person's history, including whether an individual has ever been infected and whether he or she is currently infected. These can be combined into the notion of "recovered".

As mentioned above, the finest resolution of location also includes a form of Disease State. A location's Disease State contains enough information to represent contamination. Thus, at least the "infectivity" attribute of a Location's Disease State should be maintained, although other attributes are ill-defined. Possibly, the "symptom" attribute could be used to specify whether contamination could be detected. Note that, unlike a person's Disease State, a Location will probably cycle through many infections in the course of the simulation. Whenever an infectious person is present, the Location will become contaminated. This contamination may decay quickly if the residence time in the infected state specified by the user is short.

Every Disease State except the special *dead* and *uninfected* states is associated with a probability distribution of residence times. The user may choose from a predetermined set of distributions and assign any necessary parameters.

The user may specify an arbitrary number of transitions out of each Disease State into others. Associated with each transition is a probability. Optionally, each transition may also be associated with a set of Treatment Levels. When a person leaves a Disease State, she or he will pick a new state from among those whose transitions are labeled with the person's treatment level.

There is a single consistency constraint on allowed values of attributes for a Disease State: non-zero incapacitation implies non-zero prodrome or symptoms. In particular, the following constraints are **NOT** imposed:

1. infectious => symptomatic or prodromal
2. dead => uninfectious (corpses can be hazardous)

In addition, the transition probabilities for each state must sum to unity by Treatment Level.

Some actions taken at run time during the simulation depend on thresholds set by the user. For example, when a person becomes incapacitated, he or she will skip some normal activities. Which

activities are skipped depends on the value of the person's incapacitation versus user-specified thresholds. Similarly, symptomatic people may be mis-diagnosed if their level of symptoms is not above a user-specified threshold. Also, symptomatic people may seek over-the-counter remedies or emergency care as the level of their symptoms rises. The user may specify a threshold value for *incapacitated* for staying home from any of the defined activity types. Furthermore, the user may specify any of the following thresholds for the *symptomatic* and *prodrome* attributes:

- seek over-the-counter remedies
- seek treatment at hospital/clinic
- be readily diagnosed by trained physician, lab test, etc.
- be readily diagnosed by casual observer (or contact tracer)

A dual-level threshold formulation is being implemented: when more than a user-specified number of people have been diagnosed with the disease, the second set of thresholds will be used. This allows for the increased likelihood of correct diagnosis when the disease is known to be present in the community.

Self-isolation Based on Disease Attributes

The scenario file command, *self isolate*, has been modified to include a disease state attribute and level. Multiple self-isolate commands are allowed in defining a scenario in a scenario file. The new format is as follows:

```
<time> self_isolate <disease_attribute> <disease_attribute_level>

time                    - Time as HH:MM:SS.
disease_attribute       - One of prodrome, symptoms, or incapacitation.
disease_attribute_level - Integer value of minimum disease attribute
                          value to trigger on.
```

In the EpiSimS simulation, when people go into self-isolation they go home and stay there until they no longer meet the disease attribute criteria specified. They are only following their schedule in the sense of components of time, though they continue to progress through the disease till recovery or death.

Choice of disease attribute criteria is dependent on the model of the disease and its states. One or more self-isolate scenario commands may be required to force self-isolation once people become "prodromal" till they recover or die.

Example 1:

```
00:00:00 self_isolate symptoms 2
```

Starting at time 0, people will self-isolate whenever their symptoms attribute >= 2. They will stay home until symptoms < 2.

Example 2:

```
09:00:00 self_isolate prodrome 1
09:00:00 self_isolate symptoms 1
09:00:00 self_isolate incapacitation 1
```

Starting at 9 AM on the first day, people will self-isolate whenever their prodrome attribute >= 1 or symptoms attribute >= 1 or incapacitation attribute > 1. They will continue to stay home until they no longer meet the above criteria.

Example 3:

```
10:00:00 self_isolate prodrome 2
```

Starting at 10 AM on the first day, people will self-isolate only while their prodrome attribute >= 1. Depending on the disease, this could cause people to go about their regular activities as they progress further along in the disease (if prodrome becomes 0).

Multiple Dead States Allowed in Disease Manifestation Model

Multiple dead states (where incapacitation = 10) are now allowed in a Disease Manifestation. A dead person can be infectious for a while (e.g. in a morgue) and then move to a dead state where the body is no longer infectious (ex. buried).

4.2 Implementation of Consequence Mitigation Strategies

Two kinds of treatment delivery are implemented in EpiSimS: mass delivery and ring delivery. These delivery strategies can be used to deliver any of the treatments that are specified in the disease model. Mass delivery is used to distribute treatment to the whole population at random, or targeted by demographic group. Ring delivery is targeted to those who have come into contact with a contagious person. The speed and duration of either delivery system may be limited by specified resources, such as number of treatments or number or treaters. Either delivery can be restricted to treat those within a specified range of ages. A delivery can be initiated at a prescribed time. Multiple deliveries can be carried out for a given treatment. For example, the young and the old can be vaccinated in a first round of ring delivery, while the rest can be mass vaccinated at a later time.

Prophylaxis (before infection) is modeled by changing the person's susceptibility. The user must specify a distribution of susceptibilities to use. As usual, this distribution may be conditioned on people's demographics. The variability in susceptibility post-prophylaxis allows one to model variable efficacy.

During the course of the simulation, an individual may seek treatment as described above. Availability of treatment is constrained by the simulation based on available resources (in an as-yet-to-be-determined way) and on level of symptoms (also to be determined). The simulation will determine whether each individual seeking treatment receives it, and also what level is given. Examples of possible treatment levels are:

- over-the-counter drugs
- anti-virals
- vaccinations
- antibiotics
- ventilators
- hospitalization
- morgue

The effect of treatment (after infection) is specified by the user in the Disease Manifestation Model. Each state transition may be labeled with a set of Treatment Levels. If it is not labeled, the transition is available to any individual. A labeled transition is only available to individuals who have received treatment at one of the levels included in the set.

EpiSimS has added the capability to simulate social distancing behavior or the wearing of masks that effectively reduce the infectiousness of sick persons and/or the susceptibility of susceptible persons. These behavior changes are implemented into EpiSimS through a set of new user-specified parameters: the fraction of persons by age and activity that express the behavior, and the effectiveness of the behavior in those that express it. The effectiveness of masks depends on the type of mask. The nominal mask (representing the N-95 respirator mask) is taken to reduce susceptibility by a factor of 20. Because of the one-way valve on masks designed to be worn for extended periods, this nominal mask does not reduce the infectiousness of a sick person.

4.3 Improved Disease Transmission Model

Adjustments to Transmission Rates between demographic groups is enabled by specification in the pre-processing program *InitializeHealth* (refer to Appendix B -InitializeHealth). The demographic groups may be the same as those specified for initial health states or they may differ. A transmission coefficient is assigned between every demographic group for every activity. This results in a series of tables of transmission coefficients that are used by the simulation when it calculates the transmission rate function for a particular contact. The default transmission coefficient is 1.

A transmission rate function returns the (baseline) probability of a person's becoming infected per minute of contact as a function of his/her disease transmission type and the type of an infectious person at the same location. That is, if exactly one susceptible of transmission type j and one infectious person transmission type k have been in a work location for one minute, the base probability that the susceptible has become infected is given by $\rho_{work(j,k)}$.The susceptible or infectious "person" may in fact be a location. Note that the transmission rate function need not be symmetric between susceptible and infective, and that it may be activity specific.
The reason the probability returned by the transmission rate function is called a "baseline" is that it is further adjusted by duration of contact, number of people in the location, infectivity, and susceptibility.

If more or less time than one minute has passed, the probability is adjusted as for a Poisson process, using the survival rate and assuming the probability of infection in each time interval is independent. Thus if the base probability for infection per minute is p, the probability in t time units is $1 - (1-p)^t$.

If more than one infective is present, the probability is scaled under the assumption that each infective spreads disease independently. Thus if there are N_i infectives of transmission type i, with probability of transmission p_i, the overall probability of transmission in time t will be $1 - exp\{t \Sigma_i N_i ln(1-p_i)\}$. If M susceptibles are present, we divide the probability of transmission by the scale factor M^{α}, where α is a user-specified scale factor. Each susceptible present undergoes a Bernoulli trial with the probability relevant to that person. (We'll see - if this is too computationally expensive, we could just assign the expected number of people to get infected by type.)

If the infective has infectivity r, and the susceptible has susceptibility s, the base transmission probability is adjusted to be $sr\rho_{work(j,k)}$. The user should ensure that all possible resulting

probabilities are less than unity. Taking into account the different levels of infectivity associated with each Disease State, if there are $n_{k,l}$ infectious people of type k with infectivity r, then the probability of infecting a single susceptible of type j in time t would be

$$p(t) = 1 - exp\,[t \Sigma_{types\ k} \Sigma_{infectivity\ r}\, n_{k,r}\, ln(1 - rp_{k,j})].$$

Putting everything together, the probability of infecting a person of transmission type j with susceptibility s in a location with activity type a with M susceptibles and $n_{k,l}$ infectious people of type k with infectivity r during a time t, subject to user-specified scaling in susceptibles α, is:

$$p_{j,s}(t) = \{1 - exp\,[t \Sigma_{types\ k} \Sigma_{infectivity\ r}\, n_{k,r}\, ln(1 - rs\, \rho_{a}(j,k)]\}/ M^{\alpha}$$

A Transmission Coefficients Table can be created using the Initialize Health GUI. This is a set of coefficients (<= 1) that can be specified for all combinations of selected demographics in each different building type environment (ex. home, work, shop, social recreation, serve passenger, school, and college). This can be used to increase/decrease the probability of infection based on the type of room people are in when they are exposed. A default Transmission Table of all 1's is used if one is not provided.

4.4 Sub-Location Modeling

Transims divides a city up into locations. At any moment, every person in the city is either traveling or is at a location. When a person is at a location, he is performing a specific activity, such as working or shopping.

The sublocation model in EpiSimS is used to place a person into a specific room in a specific building at the given location. Buildings have a building type, which is loosely coupled to a person's activity type. For example, if a child's activity is school, he will always be found in a school building. But if a person's activity is worker, they might be in a work building as a worker or in a school building as a teacher.

Sublocation modeling provides EpiSimS with two benefits. The first is to limit disease transmission at a location to smaller subsets of the population. The second is to allow people to go to different rooms every day, even though EpiSimS' population follows the same Groundhog Day schedule every day.

The sublocation model is also involved in finding occupants when a treater, such as a vaccinator, comes around looking for somebody to vaccinate. For example, if the vaccinator comes to a home, the vaccinator is told to vaccinate everybody in the home. If the vaccinator comes to a shopping center, nobody is found to treat. If the vaccinator comes to a school, everybody at the school is vaccinated.

4.5 New Output Events

There are now only three kinds of output events that are written to the events files during the EpiSimS simulation. One for the first time a person is exposed, another when a person is treated with

any treatment, and whenever a person changes disease state. The formats are as follows. The data on each line is tab-delimited.

Exposure event:
```
E <time> <person-id>
```

Treatment event:
```
T <time> <person-id> <household-id> <location-id> <treatment-id>
```

Change Disease event:
```
D <time> <person-id> <old-disease-state-id> <new-disease-state-id>
```

The treatment ids can be found in the treatment file. The disease state ids can be found in the disease state file. The disease state file looks like the following. The same id's are generated in EpiSimS. The disease attributes are included for use in postprocessing.

```
# This is the Disease Manifestation-State ID file.
# Each line contains the following separated by tabs.
# <ID> <manifName> <stateName> <prodrome> <symptoms> <infectivity>
<susceptibility> <incapacitated>
1       "hemorrhagic"   "dead"  0       0       0.1     0.0     10
2       "hemorrhagic"   "incubating1"   0       0       0.0     0.0     0
3       "hemorrhagic"   "infectious"    2       2       1.0     0.0     1
4       "hemorrhagic"   "sick"  2       2       1.0     0.0     4
5       "hemorrhagic"   "uninfected"    0       0       0.0     1.0     0
6       "immune"        "uninfected"    0       0       0.0     0.0     0
7       "normal, over 30"       "dead"  0       0       0.1     0.0     10
8       "normal, over 30"       "incubating1"   0       0       0.0     0.0   0
9       "normal, over 30"       "incubating2"   0       0       0.0     0.0   0
10      "normal, over 30"       "infectious"    2       0       0.1     0.0
```

4.6 Enhanced Pre- and Post-Processing

EpiSimS requires a 24 hour schedule per person in the simulation. This is built from the UPMoST activity information. The schedule generation has been enhanced to make sure everyone starts at home at 00:00:00 and ends at home at midnight. When a person's activity information does not start at 00:00:00, schedule items to start from home are added. When a person's activity information does not end at home at midnight, schedule items are added to bring them home.

The partition generation has been enhanced to scale for larger population sizes. Counts are collected for the total number of people participating in an activity for a set of aggregated locations (known as a link). Aggregation of the locations allows for more efficient sub-location modeling in the simulation.

The sample generation now collects the home location or link from the EpiSimS Demographics, due to a new organization of the UPMoST information.

The *InitializeHealth* program assigns initial health states to every person in the population. The interface allows specification of probability distributions conditioned on demographics. Time already elapsed in the initial disease state can also be specified to allow a distribution of entry times. This is most useful for disease models such as Smallpox where the residency time in various states is lengthy.

InitializeHealth is normally used to create a random set of infected people in each manifestation of the disease model. Conceivably, it can also be used to study specific time slices of a pandemic by placing the population in varying advanced stages of the disease according to previous simulation runs or research data.

The program interfaces with a database containing the UPMoST people and household entity demographics. Normal conditional query statements subdivide the population according to the disease model delineation or the particular case study. For example, the LA study and the Influenza disease model divided the population by age, whereas a former Smallpox study had a hemorrhagic subset and a female-pregnant subset.

Currently, the program considers person ID, age, gender, worker-status, household income and home location. The program is modular by design and works with any population or sample population along with any EpiSimS Disease Model. The graphic user interface enables it to be used by non-programmers. The specification of transmission coefficients is now a part of this tool. *InitializeHealth* is part of the suite of pre-processing tools that will be combined into the full Set-up Wizard.

After running a simulation, post-processing of the output events requires the merging of all the sorted output event files, processing using a number of scripts, and creation of plots. An extract script is used to create a text log and a tab-delimited count files binned by a user-supplied day fraction (ex. 1 day, 0.5 day, or 0.25 day). One of the counts files is similar to that used by the previous version for use with Gnuplot. The other contains counts for exposures, treatments, disease attributes, and each of the individual disease states that can be used by an external plotting or analysis package (e.g. Excel or Gnuplot).

Specifically counts are collected for # exposed, # ever exposed, # treated, # ever treated, # treated with X, # ever treated with X, # became, # current, and # ever for each of the disease attributes (prodromal, symptomatic, infectious, susceptible, incapacitated (includes dead), incapacitated (not dead), and dead), # became infected, # current infected, and # ever infected, # became recovered, # current recovered, # ever recovered, # disease state changes, total of disease state changes, and # became, # current, and # ever for each disease manifestation state.

Due to the large number of columns in the counts file, a user may want to select only a subset depending on what they are interested in. A few scripts have been provided to help select columns and create a user-defined subset of the counts file.

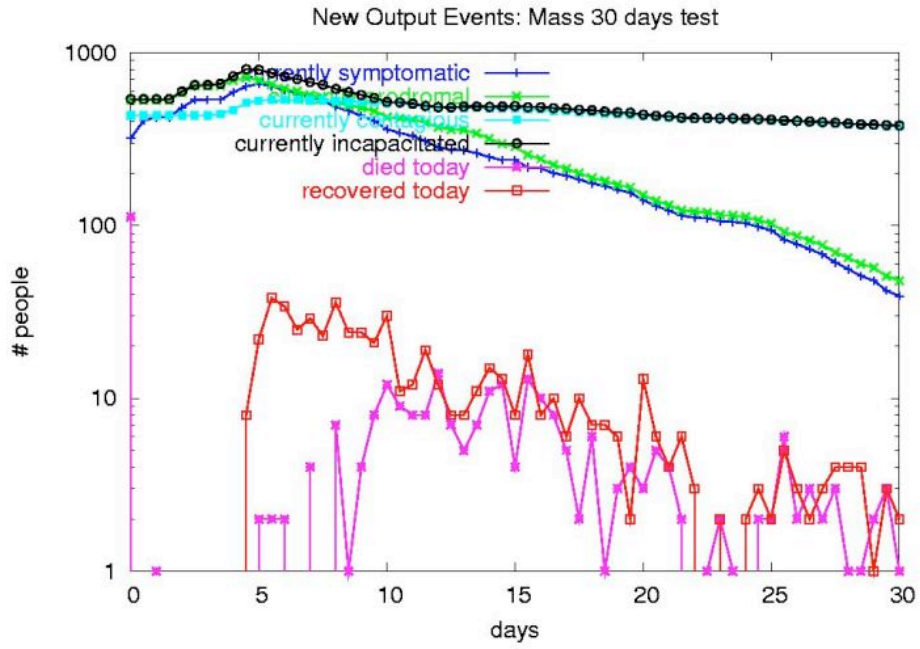


Fig. 4.6-1. Plot of output events for disease attributes using Gnuplot.

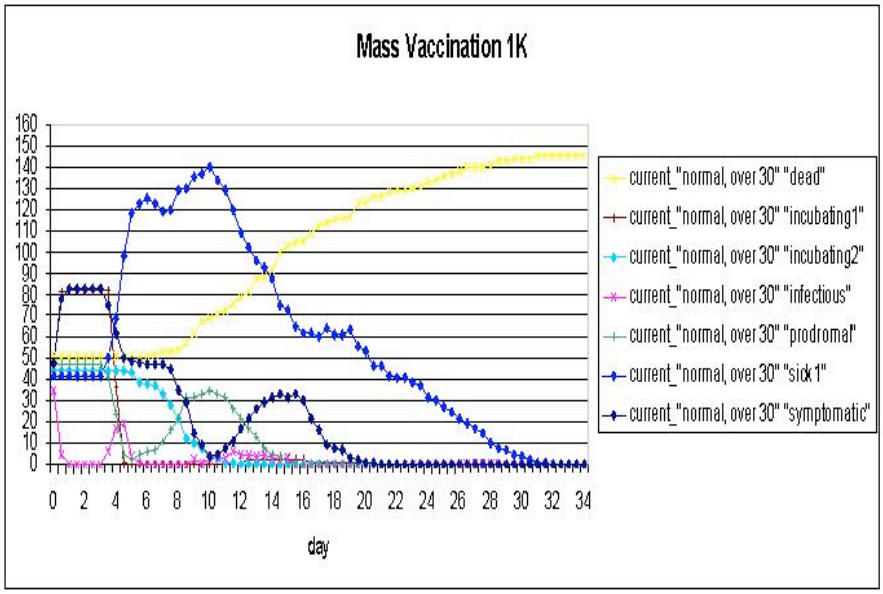


Fig. 4.6-2. Plot of output events for states of a disease manifestation in Excel.

4.7 Computational Improvements

Earliest Neighbor Parallel Synchronization Strategy

EpiSimS uses neighbor synchronization to keep the simulation in sync among the computer's processors. EpiSimS has one master process and many slave processes. Each slave has a subset of the city's locations, which it keeps for the entire simulation. As people go from place to place they are passed from one processor to another. In EpiSimS, the simulation clock time on a given processor is the time that the most recent event was popped of its event queue. Because of varying loads, some processors can get ahead of others. This mismatch then allows for the possibility that new events (e.g. arrival of individuals) are placed on an event queue with a timestamp that is earlier than the current processor simulation time. Previously, such an occurrence produced a mishandled event. A new synchronization strategy has been implemented that prevents this from occurring.

Since a person's travel time between locations is on the order of fifteen minutes, we have some latitude as to how close the time on one processor has to be to the time on another. We specify a time called the slack time, which is the amount of simulation time that a processor can simulate before they initiate synchronization with another neighbor.

When it is time for a processor to initiate synchronization, it picks another processor with which to synchronize. They compare clocks, and if one or the other is more than slack time units ahead, that one waits till the other is within slack time units of catching up. We have two methods of choosing which neighbor to choose as a synchronizing partner. One chooses the neighbor completely at random, and the other checks with the master, which selects the processor with the earliest time as the synchronizing partner.

MPI Toolbox

Due to its incompatibility with the framework software and deep hierarchical design, the MPI toolbox was rewritten to be more maintainable and understandable. The MPI toolbox is used to send and receive asynchronous messages between processors during a simulation. The toolbox provides methods for message creation, sending, receiving, and deletion. Handlers for processing the contents of a message are written in the objects using the messaging service.

Porting EpiSimS to a 64-bit Architecture

The EpiSimS code was ported to a BPROC cluster of 2048 x Intel Xeon, 2.4 GHz, 2 GB per node, Myrinet interconnect at LANL, known as Pink. LSF is used for interactive sessions (llogin) and batch jobs (bsub). Porting of the code required GCC 3.4. The external packages built were the Boost library, a new Berkeley DB library, the log4cpp library, the Metis (graph partitioning) library, and Sprng (random number generator) library. The UIS framework, UPMoST, and EpiSimS built, requiring minor changes to their build environments. UPMoST did require a minor code change due to the new Berkeley DB library. A new version of the Berkeley DB was required due to problems with the indexes created from an earlier version.

Pink nodes have less memory than our local cluster, requiring more processors for an EpiSimS simulation run. Runs are limited to 12 hours. Base case runs of the Los Angeles data of ~16M people (demographics, health, schedule, and partition files) have been run on 300 or more processors,

resulting in 160 or more simulated days. Prior to the speed-up in loading the population data, only ~111 simulated days were possible.

The schedule files must currently be read on the slave CPU's for this architecture. This constraint also caused a scaling problem with the mass delivery code, requiring some temporary recoding. This issue requires further investigation.

Memory/Scaling Improvement

A new algorithm for assigning locations to processors has reduced the computational time required to load the Los Angeles population of over 16 million persons from 4 hours to 45 minutes. This is a major improvement in efficiency that has significantly increased the number of runs that can be conducted.

The locations for the Los Angeles population activities were modeled as links. Each link represents a set of aggregated locations. This has resulted in less objects created in memory and more efficient use of the sublocation model.

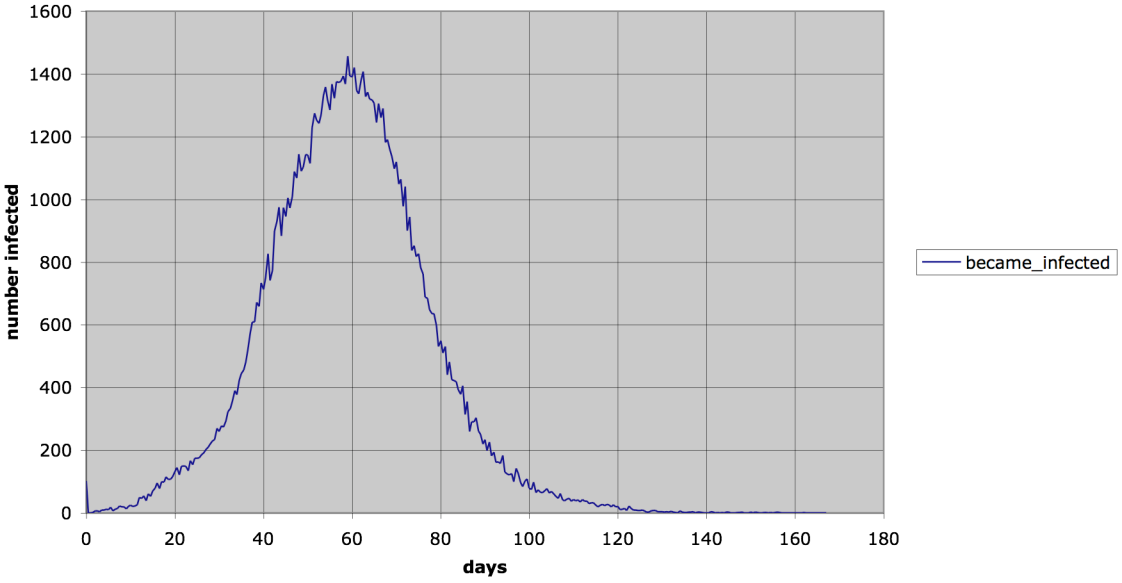
More Sophisticated GUIs

EpiSimS has added Graphical User Interfaces to do some of the simulation pre-processing. The disease model is constructed graphically using the *BuildDiseaseManifestations* tool. The initial health states are specified with the *InitializeHealth* tool, and the scenarios are defined using *ScenarioBuilder*. We are in the process of combining these tools into an *EpiSimS Set-up Suite* which will guide the user through the entire set-up process, validate the input files, and document the simulation run.

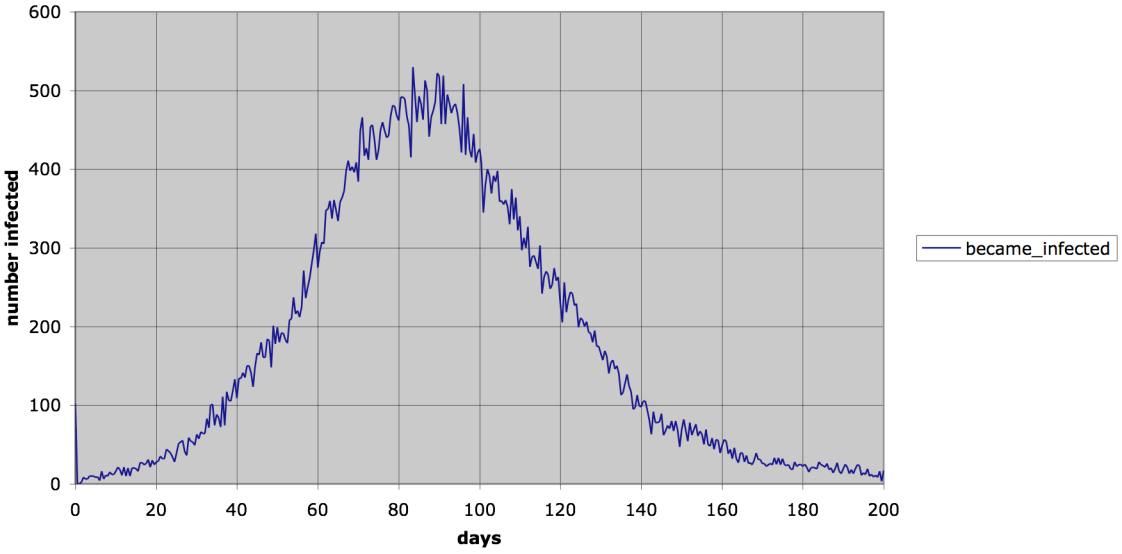
5 Verification

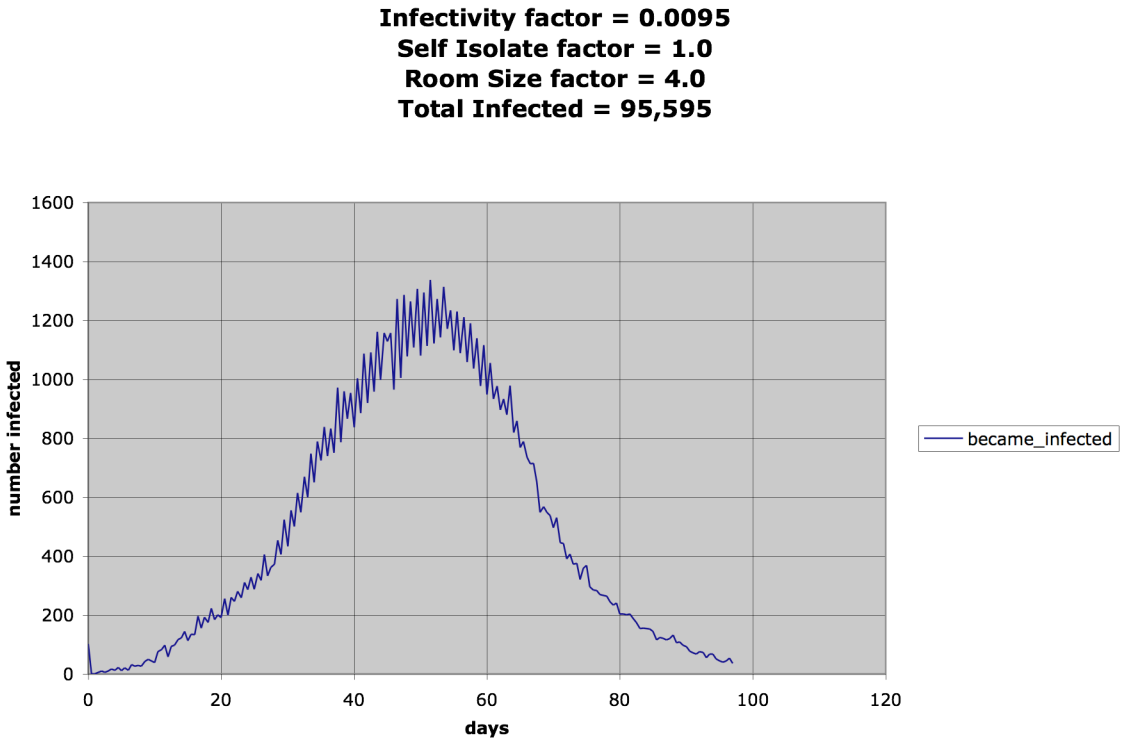
To verify that the simulation was working as it should, several input parameters were varied and the resulting simulated epidemics were analyzed. It was found to be useful to try a spectrum of variations on a smaller sample population. An EpiSimS utility creates a reduced-size population where people are more highly connected than in a random sample. The resulting population gives a good testbed to study the interdependence of EpiSim's input variables. Three variables were varied: the infectivity, the proportion of those who self isolate, and the number of people in the average room. The twelve combinations of these variations were run with the reduced-size population, and the resulting epidemic curves were examined, and found to behave as expected. The output graphs from some of these runs follows. The points on the x-axis are a half day apart, so the y-axis is the number of people who were infected in that half day period.

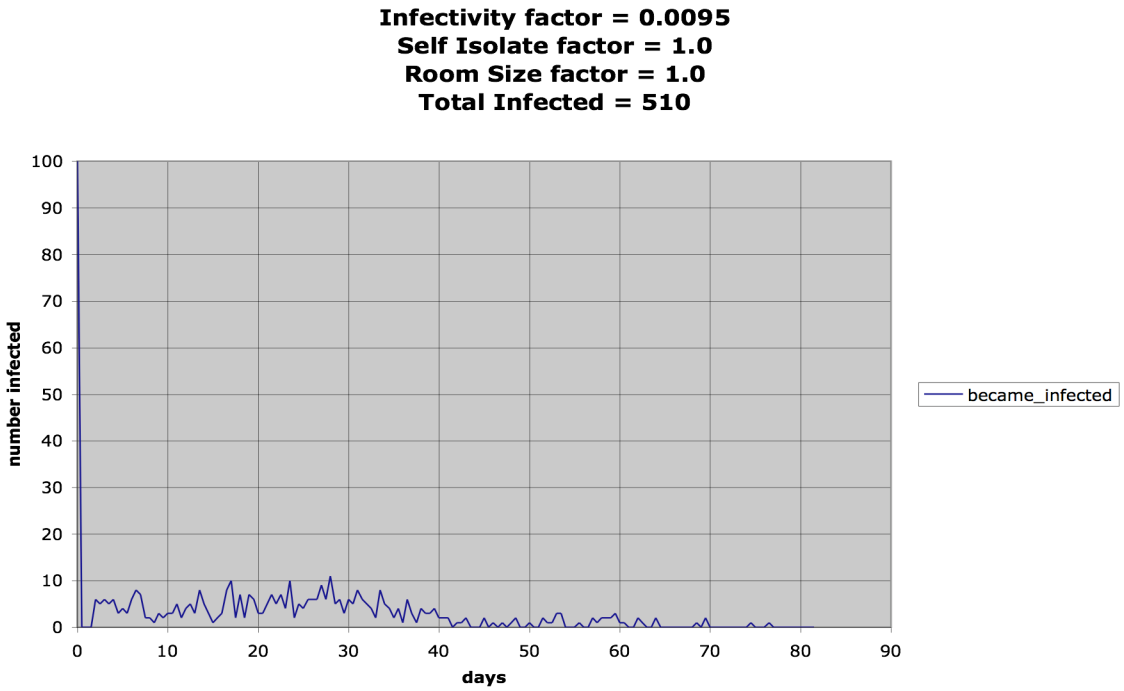
Infectivity factor = 0.015
Self Isolate factor = 0.5
Room Size factor = 1.0
Total Infected = 114,272



Infectivity factor = 0.015
Self Isolate factor = 1.0
Room Size factor = 1.0
Total Infected = 70,693







Decreasing the infectivity, increasing the number of people in a room, and decreasing the number of self isolating people all have the effect of increasing the total number of infected over the course of the epidemic, to one degree or another. They also have the effect of making the epidemic peak and die more quickly. Table 5-1 summarizes the verification runs performed on the sampled 500,000

population. The first three columns are the input parameters which we varied. The time scale represents the estimated duration of the epidemic, and the total infected is the number that were infected during the course of the run.

Infectivity	Self-isolate	Roomsize	Time scale	Total infected
0.0095	0.5	0.5	3.5	246
0.0095	0.5	1	90.5	52870
0.0095	0.5	4	36.5	118112
0.0095	1	0.5	0	182
0.0095	1	1	19.5	510
0.0095	1	4	50.5	95595
0.0150	0.5	0.5	103	55536
0.0150	0.5	1	59	114272
0.0150	0.5	4	31	172033
0.0150	1	0.5	4	254
0.0150	1	1	88.5	70693
0.0150	1	4	36	142781

Table 5-1. Verification test results on reduced-size population. The nominal infectiousness in transmissions per hour per infectious person is obtained by multiplying the infectivity multiplier value shown by 60*0.005.

6 Results: Epidemic Simulation Without Disease Interventions

6.1 Epidemiologic Results

The Epidemic Curve

The base case scenario for an avian influenza epidemic in Los Angeles has 16,106,535 residents, each represented as an individual in the EpiSimS simulation. Each person has a schedule of activities that they undertake throughout the day. There are eight types of activity: *home, work, shopping, visiting, social recreation, service passengers (e.g. drive a carpool), school*, and *college*; plus a ninth activity designated *other*. Each activity occurs in a room, and there may be other people in the room at the same time. The virus can be transmitted between persons that occupy the same room at the same time. There are 562,452 Los Angeles *locations* represented in the EpiSimS simulation. At each location, there is one *building* for each of the eight designated activity categories. Each building contains rooms: the number of rooms for each activity type at each location is derived from the Los Angeles data. For some activity categories, e.g. home or work, a person does that activity in the same room each day, with the same other people. For other activity categories, such as shopping, the person is assigned to a randomly selected room in the building corresponding to that activity category, at the location designated in her schedule. The activity schedules and the locations are statistically the same as those of actual people, and are constructed from a variety of real data sources. In addition, an urban mobility simulation computes the travel time between activities, accounting for distance, roads and traffic.

The base case treats the scenario that an avian strain makes so rapid a jump to humans that no vaccine can be developed in time, and that nobody has developed immunity to the disease. Further, in the base case, it is taken that there are no effective antiviral treatments or vaccines are available.

Some persons who get infected will discontinue their activity schedule: when they get sick, they go home and remain there through the course of their symptoms. 33% of adults and seniors will self-isolate in this fashion. 50% of students will self-isolate, as will 53.3% of pre-school children. Persons are assigned to self-isolate at random, based on their age and the age-dependent probability of self-isolating. Those persons that self-isolate can infect members of their household, but will no longer spread the virus outside their household.

The target attack rate for the base case was taken to be 25%, for consistency with the WHO baseline avian flu pandemic scenario. The infectiousness was scaled to obtain a pandemic that infected about 4 million of the individuals in the simulation. The calibration and infectiousness scaling is described in section 3.3.

The base case run starts with the infection (at time = 0.0 days) of 202 randomly selected individuals. This represents a compromise between having enough infected individuals at the start to average the fluctuations in transmission histories, and being able to compute the trajectory beginning early in the outbreak.

The base case EpiSimS simulation was run from July 16, 2005 to July 23, 2005, distributed over 106 processors, each with 2 GB of local memory. This run of the EpiSimS simulation is designated *try30*. The number of new infections per day is shown in Fig. 6.1-1, and as a logarithm in Fig. 6.1-2.

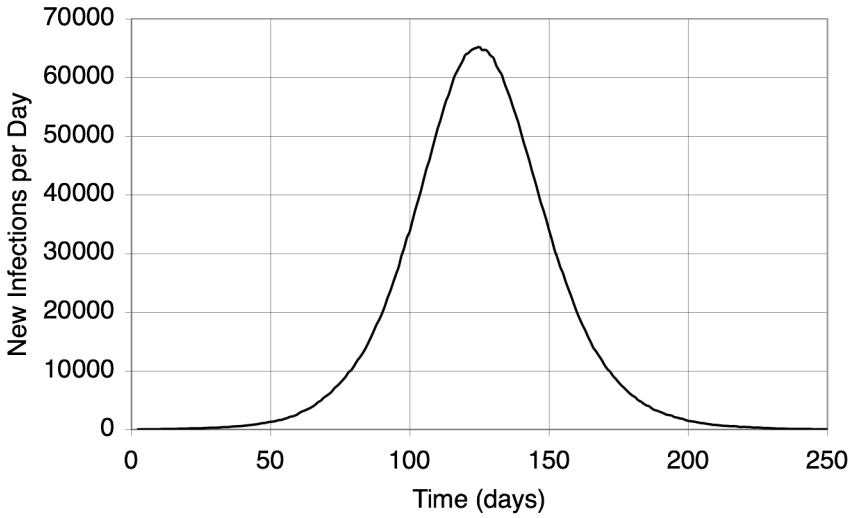


Fig. 6.1-1. The base case EpiSimS simulation run, *try30*, showing the number of new infections per day, for pandemic influenza in a population of 16.1 million individuals.

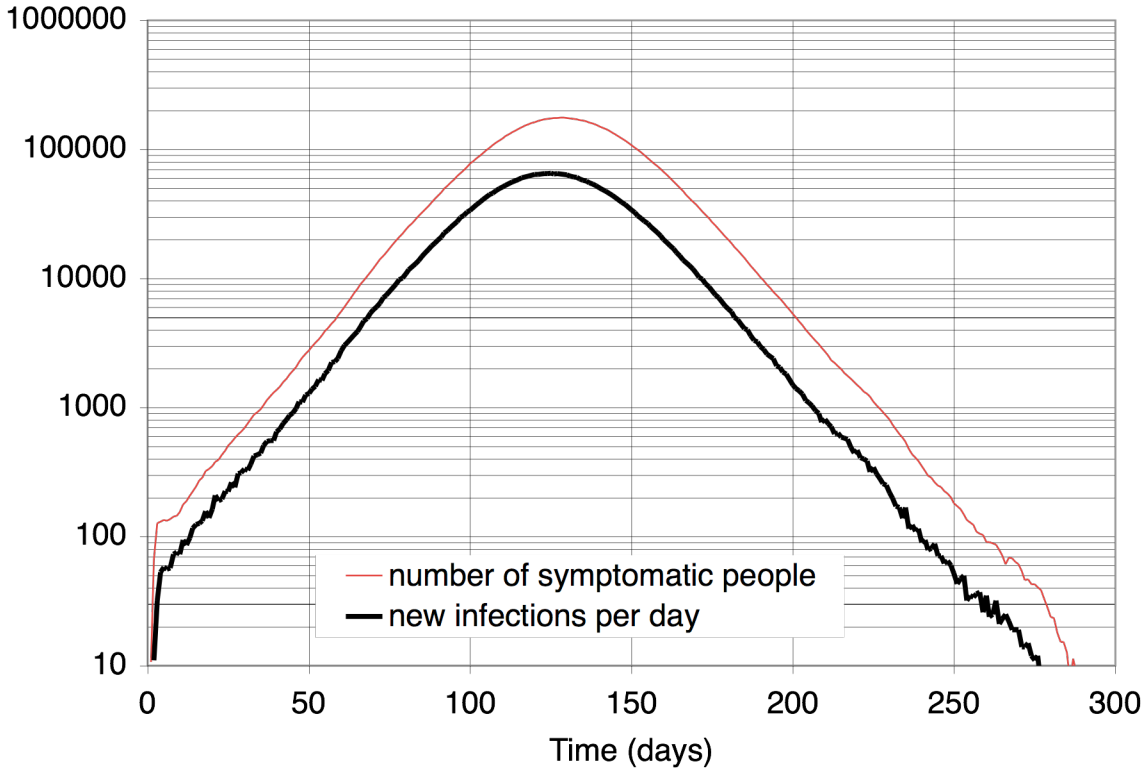


Fig. 6.1-2. The base case EpiSimS simulation run of pandemic influenza in a Los Angeles population of 16.1 million individuals, *try30*, showing the number of new cases per day and the current number of symptomatic persons.

The cumulative number of cases is shown in Fig. 6.1-3.

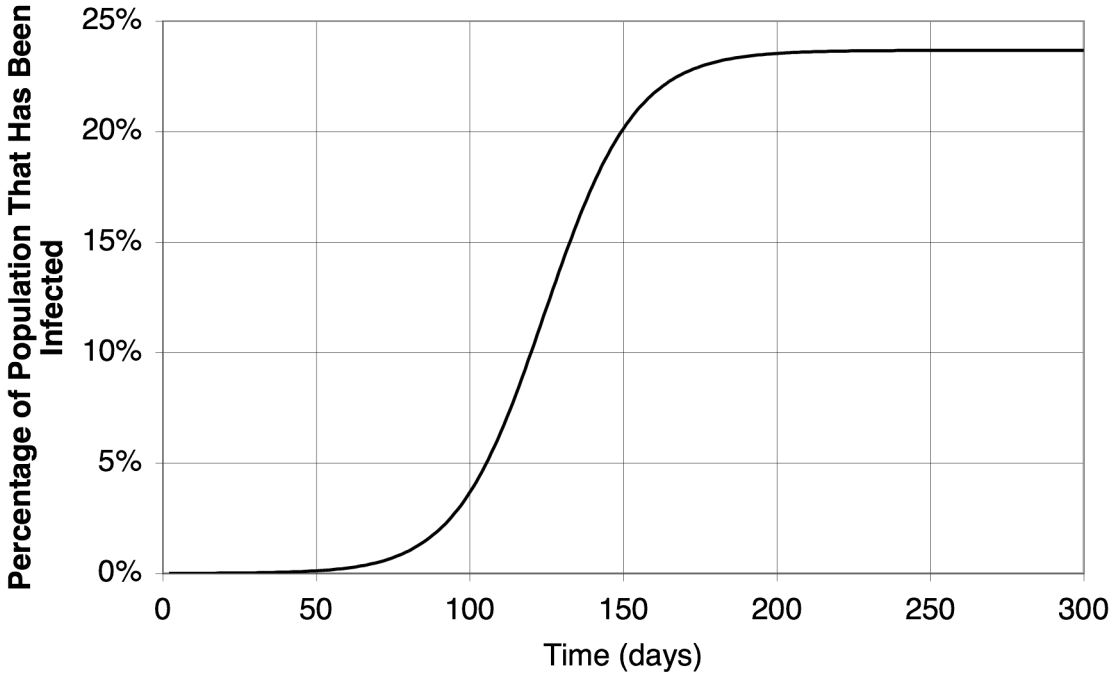


Fig. 6.1-3. The base case EpiSimS simulation run, *try30*, showing the cumulative number of cases, as a percentage of the initial population.

The peak new case rate of 65,278 new cases per day occurs from 124.0 to 125.0 days after the start of the epidemic. By day 125, when the peak new case rate occurs, the cumulative number of infections reaches 1,938,657. The simulation was run out to day 307, by which time 3,813,957 people had been infected and no new infections occur.

Although the "duration" of the epidemic is a somewhat nebulous notion, it can be quantified as follows. The new infection rate exceeds half its peak value for a total of 52 days, from 26 days before the peak until 25 days after the peak. Alternatively, the middle 80% of all cases occur during a period of 64 days, again roughly centered around the 50% point.

The epidemic curves normalized to the population to give results that are essentially independent of the population. As a percentage of the initial population, the base case reaches a peak new infection rate of 0.405% of the initial population becoming infected per day. The percentage of the population that is currently symptomatic is shown in Fig. 6.1-4. This number translates to hospital/clinic visits, hospitalization rates, and fatalities.

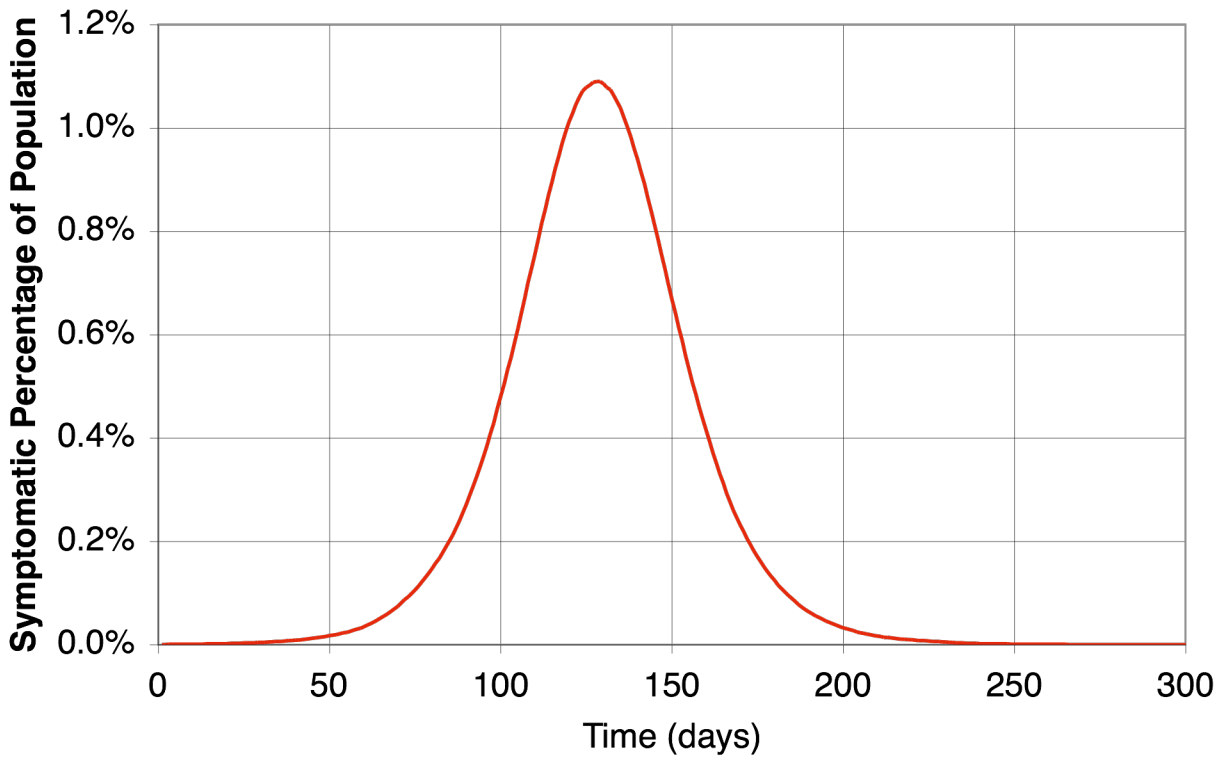


Fig. 6.1-4. The fraction of the population that is symptomatic, in the base case pandemic influenza simulation of Los Angeles.

What are people doing when they become infected?

The simulation keeps a record of the activity that each person was doing when he became infected. Fig. 6.1-5 shows the breakout of new cases per day into the activity categories. The breakout of cumulative infections by activity category for the whole epidemic is shown in Table 6.1-1. More infections are acquired at home than either school or work, and more infections are acquired at work than at school. However, the infections acquired at school per student exceeds the infections acquired at work per worker, leading to a higher attack rate among students than among workers.

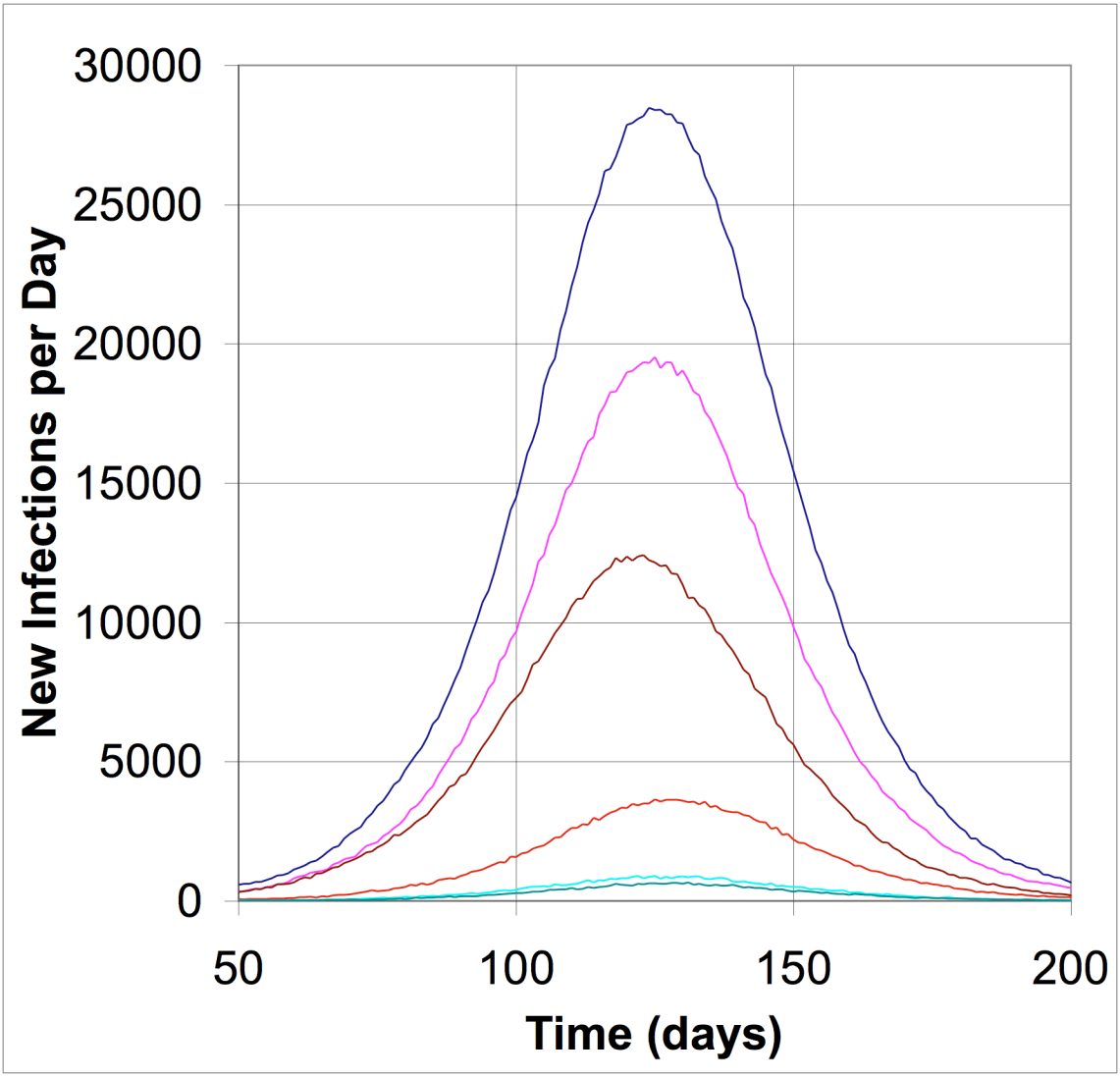


Fig. 6.1-5. Breakout of new cases per day into categories of where the person was when she was infected. In descending order, the curves are Home, Work, School, Shop, Social Recreation Building, College.

Where infection is acquired	Fraction of cumulative infections
Home	**43.9%**
Work	**29.2%**
School (K-12)	**18.9%**
Shop	**5.6%**
Social recreation	**1.3%**
College	**1.0%**

Table 6.1-1. Breakout of infections by activity category

Fig. 6.1-6 shows the same data as that in Fig. 6.1-5, except that the number of new cases per day acquired in each activity category is normalized to the total daily number of new cases. This highlights the timing: the pandemic runs through schools earlier than through the rest of the population. Only the home, work, and school categories are shown.

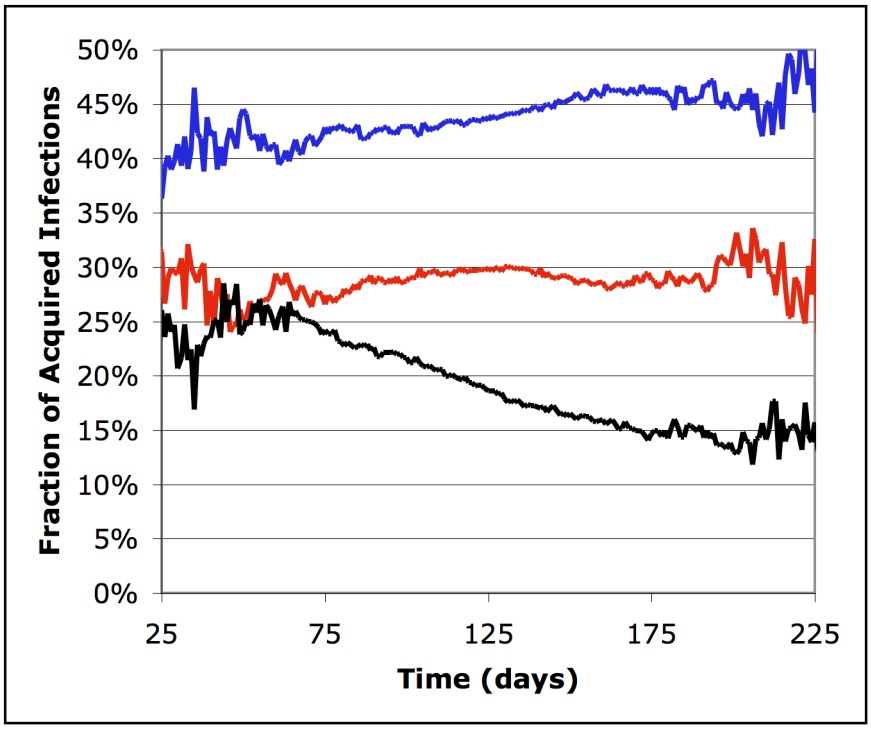


Fig. 6.1-6. Fraction of daily new cases that are acquired (curves in descending order) at Home, Work, and School, illustrating the early wave of school-acquired infections.

Age dependence of infections

The EpiSimS simulation generates records of how the epidemic depends on people's age. Fig. 6.1-7 shows the distribution of ages in the initial population of Los Angeles. Fig. 6.1-8 shows the current fraction of people of each age that are sick at three times: on day 97, five days after the peak on day 130, and well after the peak on day 160. The relative rise in the fraction of school-age persons shows the early onset of a wave of school-acquired infections.

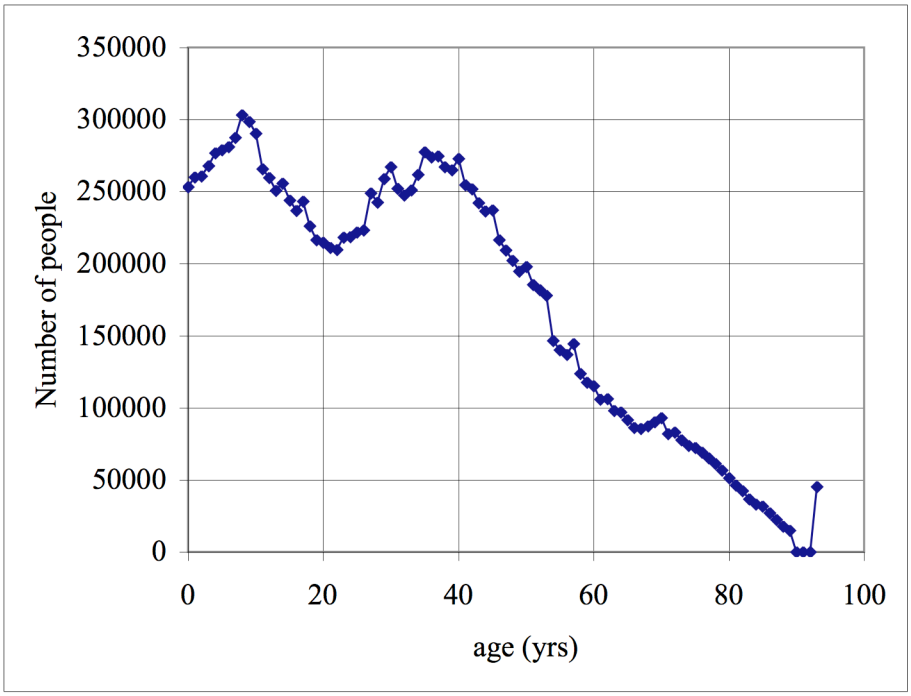


Fig. 6.1-7. The distribution of ages in the Los Angeles population

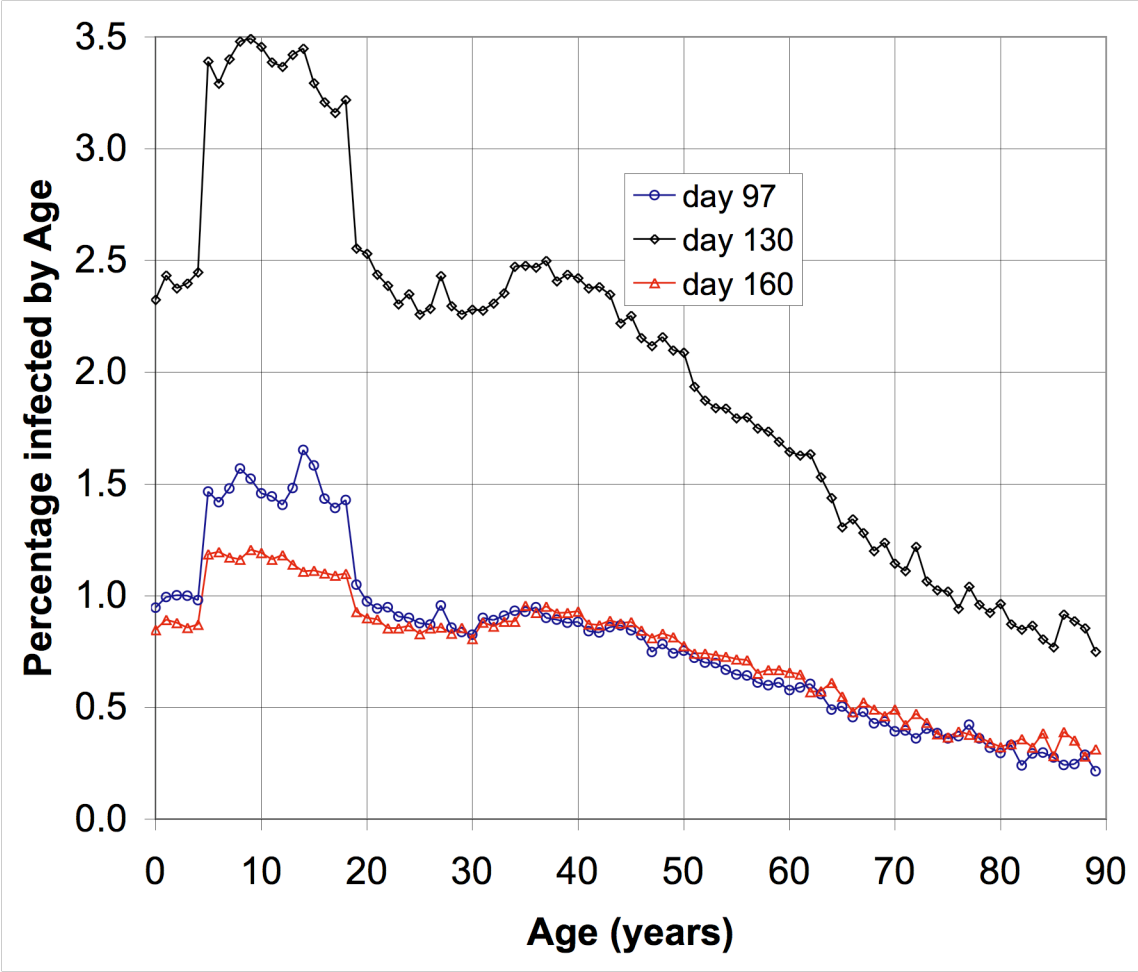


Fig. 6.1-8. The fraction of persons of various ages that are infected before the peak of the outbreak on day 97, five days after the peak on day 130, and well after the peak on day 160.

Other measures of epidemic trajectory

In addition to new cases per day, EpiSimS can generate the number of people in the incubating stage, the number of people in the symptomatic stage, the number of remaining susceptible people, etc. A list of 226 fields of data extracted from the base case runs is shown in Table 6.1-2.

exposed	became_"preschool" "treated_sub-clinical_infectious1"
ever_exposed	current_"preschool" "treated_sub-clinical_infectious1"
treated	ever_"preschool" "treated_sub-clinical_infectious1"
ever_treated	became_"preschool" "treated_sub-clinical_infectious2"
treated_"antivirals"	current_"preschool" "treated_sub-clinical_infectious2"
ever_treated_"antivirals"	ever_"preschool" "treated_sub-clinical_infectious2"
treated_"vaccine"	became_"preschool" "treated_symptomatic_circulating1"
ever_treated_"vaccine"	current_"preschool" "treated_symptomatic_circulating1"
treated_"antivirals,vaccine"	ever_"preschool" "treated_symptomatic_circulating1"
ever_treated_"antivirals,vaccine"	became_"preschool" "treated_symptomatic_circulating2"
became_prodromal	current_"preschool" "treated_symptomatic_circulating2"

current_prodromal
ever_prodromal
became_symptomatic
current_symptomatic
ever_symptomatic
became_infectious
current_infectious
ever_infectious
became_susceptible
current_susceptible
ever_susceptible
became_incapacitated (includes dead)
current_incapacitated (includes dead)
ever_incapacitated (includes dead)
became_incapacitated (not dead)
current_incapacitated (not dead)
ever_incapacitated (not dead)
became_dead
current_dead
ever_dead
became_infected
current_infected
ever_infected
became_recovered
current_recovered
ever_recovered
never_infected
state_changes
cum_state_changes
became_"adult" "dead"
current_"adult" "dead"
ever_"adult" "dead"
became_"adult" "latent_incubating"
current_"adult" "latent_incubating"
ever_"adult" "latent_incubating"
became_"adult" "sub-clinical_infectious"
current_"adult" "sub-clinical_infectious"
ever_"adult" "sub-clinical_infectious"
became_"adult" "symptomatic_circulating"
current_"adult" "symptomatic_circulating"
ever_"adult" "symptomatic_circulating"
became_"adult" "symptomatic_non-circulating"
current_"adult" "symptomatic_non-circulating"
ever_"adult" "symptomatic_non-circulating"
became_"adult" "treatable_sub-clinical_infectious"
current_"adult" "treatable_sub-clinical_infectious"
ever_"adult" "treatable_sub-clinical_infectious"
became_"adult" "treatable_symptomatic_circulating"
current_"adult" "treatable_symptomatic_circulating"
ever_"adult" "treatable_symptomatic_circulating"
became_"adult" "treatable_symptomatic_non-circulating"
current_"adult" "treatable_symptomatic_non-circulating"
ever_"adult" "treatable_symptomatic_non-circulating"
became_"adult" "treated_sub-clinical_infectious1"
current_"adult" "treated_sub-clinical_infectious1"
ever_"adult" "treated_sub-clinical_infectious1"
became_"adult" "treated_sub-clinical_infectious2"
current_"adult" "treated_sub-clinical_infectious2"
ever_"adult" "treated_sub-clinical_infectious2"
became_"adult" "treated_symptomatic_circulating1"
current_"adult" "treated_symptomatic_circulating1"
ever_"adult" "treated_symptomatic_circulating1"
became_"preschool" "dead"
current_"preschool" "dead"
ever_"preschool" "dead"
became_"preschool" "latent_incubating"
current_"preschool" "latent_incubating"
ever_"preschool" "latent_incubating"
became_"preschool" "sub-clinical_infectious"
current_"preschool" "sub-clinical_infectious"
ever_"preschool" "sub-clinical_infectious"
became_"preschool" "symptomatic_circulating"
current_"preschool" "symptomatic_circulating"
ever_"preschool" "symptomatic_circulating"
became_"preschool" "symptomatic_non-circulating"
current_"preschool" "symptomatic_non-circulating"
ever_"preschool" "symptomatic_non-circulating"
became_"preschool" "treatable_sub-clinical_infectious"
current_"preschool" "treatable_sub-clinical_infectious"
ever_"preschool" "treatable_sub-clinical_infectious"
became_"preschool" "treatable_symptomatic_circulating"
current_"preschool" "treatable_symptomatic_circulating"
ever_"preschool" "treatable_symptomatic_circulating"
became_"preschool" "treatable_symptomatic_non-circulating"
current_"preschool" "treatable_symptomatic_non-circulating"
ever_"preschool" "treatable_symptomatic_non-circulating"
became_"preschool" "treated_sub-clinical_infectious1"
current_"preschool" "treated_sub-clinical_infectious1"
ever_"preschool" "treated_sub-clinical_infectious1"
became_"preschool" "treated_sub-clinical_infectious2"
current_"preschool" "treated_sub-clinical_infectious2"
ever_"preschool" "treated_sub-clinical_infectious2"
became_"preschool" "treated_symptomatic_circulating1"
current_"preschool" "treated_symptomatic_circulating1"
ever_"preschool" "treated_symptomatic_circulating1"
became_"preschool" "treated_symptomatic_circulating2"
current_"preschool" "treated_symptomatic_circulating2"
ever_"preschool" "treated_symptomatic_circulating2"
became_"preschool" "treated_symptomatic_non-circulating1"
current_"preschool" "treated_symptomatic_non-circulating1"
ever_"preschool" "treated_symptomatic_non-circulating1"
became_"preschool" "treated_symptomatic_non-circulating2"
current_"preschool" "treated_symptomatic_non-circulating2"
ever_"preschool" "treated_symptomatic_non-circulating2"
became_"preschool" "uninfected"
current_"preschool" "uninfected"
ever_"preschool" "uninfected"
became_"room" "uninfected"
current_"room" "uninfected"
ever_"room" "uninfected"
became_"senior" "dead"
current_"senior" "dead"
ever_"senior" "dead"
became_"senior" "latent_incubating"
current_"senior" "latent_incubating"
ever_"senior" "latent_incubating"
became_"senior" "sub-clinical_infectious"
current_"senior" "sub-clinical_infectious"
ever_"senior" "sub-clinical_infectious"
became_"senior" "symptomatic_circulating"
current_"senior" "symptomatic_circulating"
ever_"senior" "symptomatic_circulating"
became_"senior" "symptomatic_non-circulating"
current_"senior" "symptomatic_non-circulating"
ever_"senior" "symptomatic_non-circulating"
became_"senior" "treatable_sub-clinical_infectious"
current_"senior" "treatable_sub-clinical_infectious"
ever_"senior" "treatable_sub-clinical_infectious"
became_"senior" "treatable_symptomatic_circulating"
current_"senior" "treatable_symptomatic_circulating"
ever_"senior" "treatable_symptomatic_circulating"
became_"senior" "treatable_symptomatic_non-circulating"
current_"senior" "treatable_symptomatic_non-circulating"
ever_"senior" "treatable_symptomatic_non-circulating"
became_"senior" "treated_sub-clinical_infectious1"
current_"senior" "treated_sub-clinical_infectious1"
ever_"senior" "treated_sub-clinical_infectious1"
became_"senior" "treated_sub-clinical_infectious2"
current_"senior" "treated_sub-clinical_infectious2"
ever_"senior" "treated_sub-clinical_infectious2"
became_"senior" "treated_symptomatic_circulating1"
current_"senior" "treated_symptomatic_circulating1"
ever_"senior" "treated_symptomatic_circulating1"
became_"senior" "treated_symptomatic_circulating2"
current_"senior" "treated_symptomatic_circulating2"
ever_"senior" "treated_symptomatic_circulating2"
became_"senior" "treated_symptomatic_non-circulating1"
current_"senior" "treated_symptomatic_non-circulating1"
ever_"senior" "treated_symptomatic_non-circulating1"
became_"senior" "treated_symptomatic_non-circulating2"
current_"senior" "treated_symptomatic_non-circulating2"
ever_"senior" "treated_symptomatic_non-circulating2"

became_"adult" "treated_symptomatic_circulating2"	became_"senior" "uninfected"
current_"adult" "treated_symptomatic_circulating2"	current_"senior" "uninfected"
ever_"adult" "treated_symptomatic_circulating2"	ever_"senior" "uninfected"
became_"adult" "treated_symptomatic_non-circulating1"	became_"youth" "dead"
current_"adult" "treated_symptomatic_non-circulating1"	current_"youth" "dead"
ever_"adult" "treated_symptomatic_non-circulating1"	ever_"youth" "dead"
became_"adult" "treated_symptomatic_non-circulating2"	became_"youth" "latent_incubating"
current_"adult" "treated_symptomatic_non-circulating2"	current_"youth" "latent_incubating"
ever_"adult" "treated_symptomatic_non-circulating2"	ever_"youth" "latent_incubating"
became_"adult" "uninfected"	became_"youth" "sub-clinical_infectious"
current_"adult" "uninfected"	current_"youth" "sub-clinical_infectious"
ever_"adult" "uninfected"	ever_"youth" "sub-clinical_infectious"
became_"immune" "uninfected"	became_"youth" "symptomatic_circulating"
current_"immune" "uninfected"	current_"youth" "symptomatic_circulating"
ever_"immune" "uninfected"	ever_"youth" "symptomatic_circulating"
became_"preschool" "dead"	became_"youth" "symptomatic_non-circulating"
current_"preschool" "dead"	current_"youth" "symptomatic_non-circulating"
ever_"preschool" "dead"	ever_"youth" "symptomatic_non-circulating"
became_"preschool" "latent_incubating"	became_"youth" "treatable_sub-clinical_infectious"
current_"preschool" "latent_incubating"	current_"youth" "treatable_sub-clinical_infectious"
ever_"preschool" "latent_incubating"	ever_"youth" "treatable_sub-clinical_infectious"
became_"preschool" "sub-clinical_infectious"	became_"youth" "treatable_symptomatic_circulating"
current_"preschool" "sub-clinical_infectious"	current_"youth" "treatable_symptomatic_circulating"
ever_"preschool" "sub-clinical_infectious"	ever_"youth" "treatable_symptomatic_circulating"
became_"preschool" "symptomatic_circulating"	became_"youth" "treatable_symptomatic_non-circulating"
current_"preschool" "symptomatic_circulating"	current_"youth" "treatable_symptomatic_non-circulating"
ever_"preschool" "symptomatic_circulating"	ever_"youth" "treatable_symptomatic_non-circulating"
became_"preschool" "symptomatic_non-circulating"	became_"youth" "treated_sub-clinical_infectious1"
current_"preschool" "symptomatic_non-circulating"	current_"youth" "treated_sub-clinical_infectious1"
ever_"preschool" "symptomatic_non-circulating"	ever_"youth" "treated_sub-clinical_infectious1"
became_"preschool" "treatable_sub-clinical_infectious"	became_"youth" "treated_sub-clinical_infectious2"
current_"preschool" "treatable_sub-clinical_infectious"	current_"youth" "treated_sub-clinical_infectious2"
ever_"preschool" "treatable_sub-clinical_infectious"	ever_"youth" "treated_sub-clinical_infectious2"
became_"preschool" "treatable_symptomatic_circulating"	became_"youth" "treated_symptomatic_circulating1"
current_"preschool" "treatable_symptomatic_circulating"	current_"youth" "treated_symptomatic_circulating1"
ever_"preschool" "treatable_symptomatic_circulating"	ever_"youth" "treated_symptomatic_circulating1"
became_"preschool" "treatable_symptomatic_non-circulating"	became_"youth" "treated_symptomatic_circulating2"
current_"preschool" "treatable_symptomatic_non-circulating"	current_"youth" "treated_symptomatic_circulating2"
ever_"preschool" "treatable_symptomatic_non-circulating"	ever_"youth" "treated_symptomatic_circulating2"
	became_"youth" "treated_symptomatic_non-circulating1"
	current_"youth" "treated_symptomatic_non-circulating1"
	ever_"youth" "treated_symptomatic_non-circulating1"
	became_"youth" "treated_symptomatic_non-circulating2"
	current_"youth" "treated_symptomatic_non-circulating2"
	ever_"youth" "treated_symptomatic_non-circulating2"
	became_"youth" "uninfected"
	current_"youth" "uninfected"
	ever_"youth" "uninfected"

Table 6.1-2. The data fields used to extract information of interest from EpiSimS simulation runs

Fig. 6.1-9 shows an example of the trajectories of current counts of people in various stages of illness for the base case EpiSimS run. The current numbers of people in the incubation stage and the infectious stage are shown. The maximum number of infectious people occurs 2 days after the maximum number of incubating people. Some of the infectious people are symptomatic, and some

of the symptomatic people are incapacitated. These are both shown of Fig. 6.1-9, as is the number of people that are currently dead from the disease.

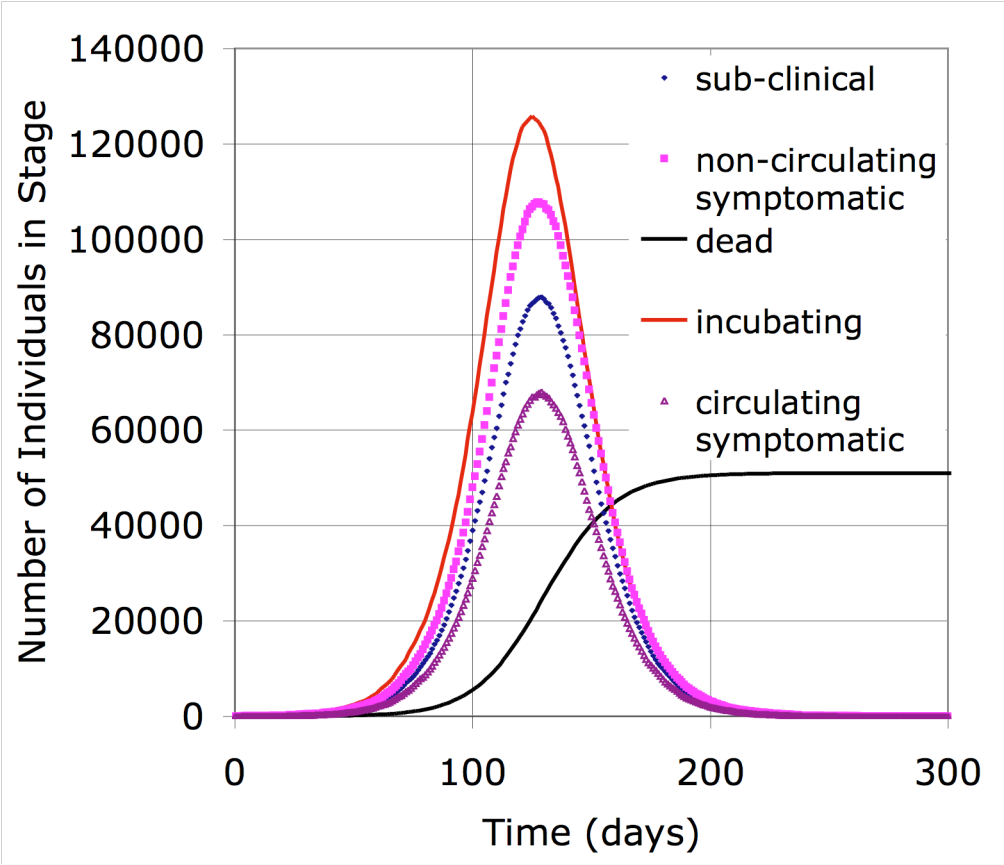


Fig. 6.1-9. Current numbers of people in various disease stages, for base case run3080. The number of people in the infectious stage is further broken out into those that are symptomatic, and those that are incapacitated

6.2 Observation of Power-Law Mixing

Fig. 6.2-1 shows the number of new cases per day per infectious person, as a function of the fraction of the initial population that is still susceptible, as generated by the *try30* base case EpiSimS simulation run.

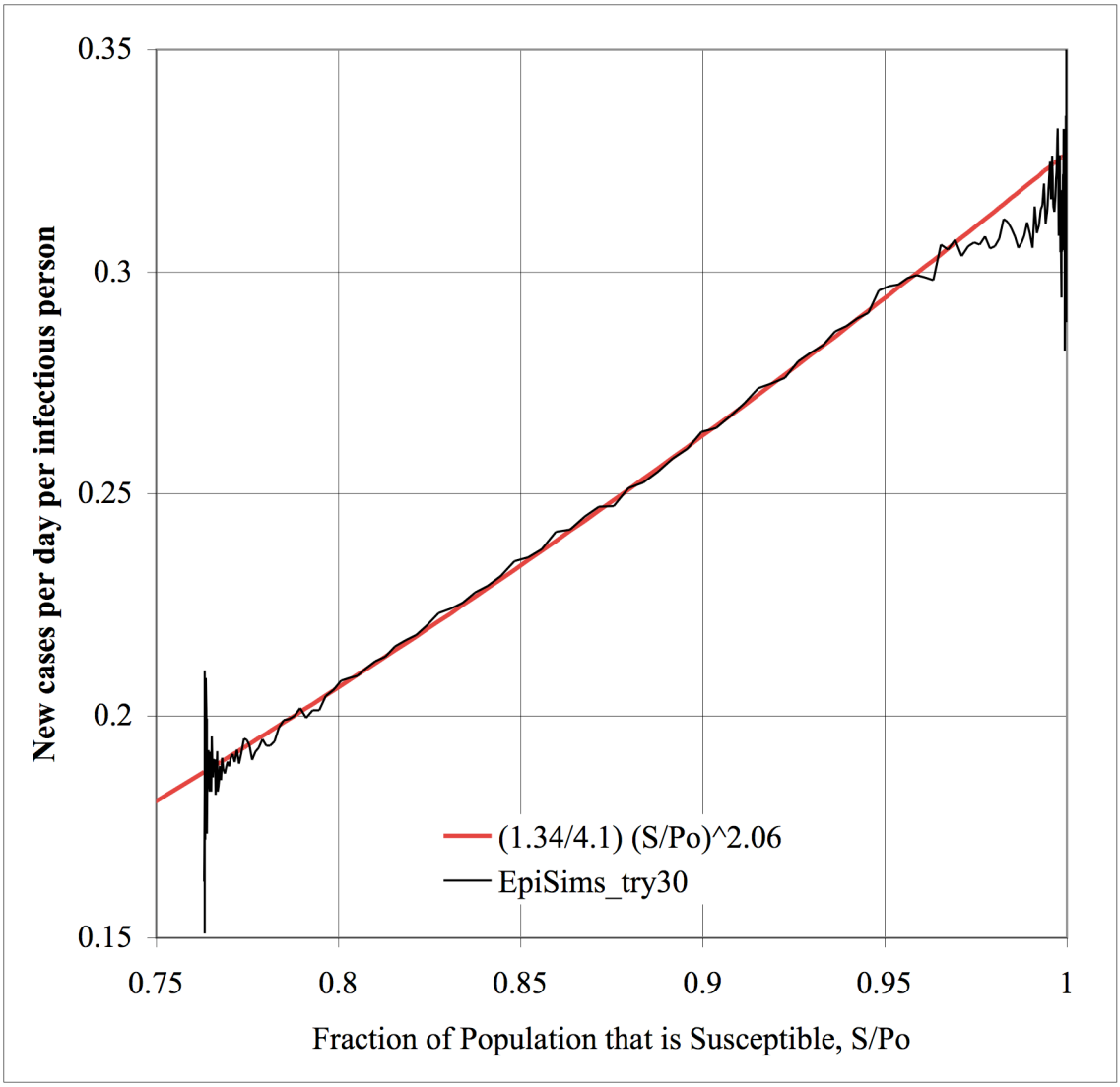


Fig. 6.2-1. The number of new cases per day per infectious person, as a function of the fraction of the initial population that is still susceptible, from EpiSimS base case run *try30*.

Traditional epidemiology modeling takes the number of new cases per day, q to be the number of infectious people, *I*, times the basic number of transmissions per day per infectious person, times a correction to account for reduction in the susceptible fraction:

$$q = I(R_0 / \tau_I)(S / P_0)$$

R_0 is the basic reproductive number, giving the average number of infections that each sick person would transmit if the entire population was susceptible, and τ_I is the average duration of the infectious stage. The ratio (R_0/τ_I) is thus the average number of transmissions per day per infectious person for a completely susceptible population. Alternatively, the basic number of transmissions per day per infectious person can be formulated as a weighted sum of terms formed as products of the

number of contacts per person, the hours per contact, and the transmission probability per contact hour. The final factor implements the common-sense notion that the transmission rate ought to be proportional to the fraction of one's contacts that are susceptible, which ought to be the fraction of the whole population that is susceptible.

The base case EpiSimS run, however shows that for the social contact structure connecting the people of Los Angeles, the number of transmissions per infectious person falls off with dropping susceptible fraction at much higher than a linear scaling. As can be seen in Fig. 6.2-1, the number of new cases per day per infectious person is approximated by 0.328 times the susceptible fraction of the original population raised to the 2.06 power. The number of new cases per day is expressed as

$$q_{LA} = I(R_0 / \tau_I)(S_0 / P_0)(S / S_0)^{2.06}$$

The reason for this power-law scaling is that the early infections occur disproportionally higher in persons with larger number of contacts. Then, as the epidemic progresses, the remaining susceptible people have on average fewer and fewer contacts, as do the infectious persons. The exponent of 2.06 is within the range of the power-law exponent describing the network connectivity of several large social networks [Albert 2003, Barabasi 2004].

Taking the average duration of the infectious stage to be 4.1 days, the coefficient extracted from the try30 EpiSimS run, i.e. 0.328, indicates a basic R_0 value of 1.34 transmissions per case for the simulated epidemic.

The Los Angeles power-law new-case-rate scaling has been implemented into the scoping model. Fig. 6.2-2 shows a comparison between the half-day-bucket-transition model and the EpiSimS simulation of the base case, try30.

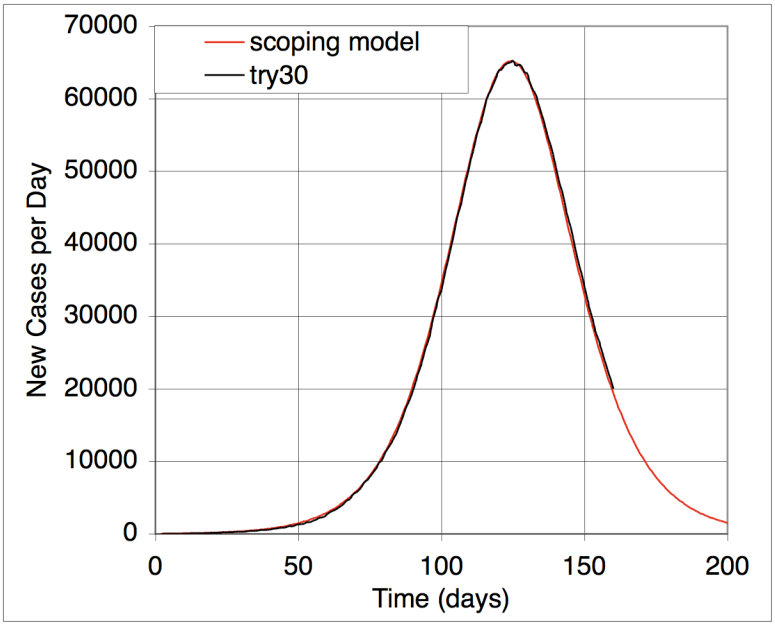


Fig. 6.2-2. Comparison of new cases per day, for an EpiSimS simulation of the base case, and the half-day-bucket-transition model.

Relative to the EpiSimS base case simulation *try30*, the scoping model used 1) the same initial population of 16,106,535 people, 2) the same 202 initial index cases infected at the start of the epidemic, 3) an initial infection rate of 1.34 transmissions per index case, equivalent to 0.327 transmissions per day per infectious person, 4) the scaling of the transmissions per infectious person as the susceptible fraction raised to the 2.1 power, and of course 5) the same incubation and infectious stage duration histograms. The scoping model can reproduce the epidemic curve generated by the EpiSimS simulation given the scaling exponent and coefficient of the transmissions per infectious person expression. However, the EpiSimS simulation is the only method available to determine the scaling exponent and coefficient for the data-based social contact network of a real city.

An epidemiology model (using the reduction in transmissions per case as proportional to the susceptible fraction) will not be able to reproduce the epidemic curve. Even by imposing an unjustifiable reduced "effective population" and adjusting the reproductive number, the best obtainable epidemic curves do not match the EpiSimS results nearly as well as the power-law scaling.

6.3 Geospatial Epidemic Dynamics

The EpiSimS epidemic simulation has a unique capability for geospatial visualization of epidemic dynamics. The progression of the base case epidemic is illustrated in Figs. 6.3-1 through 6.3-3, showing the state of infection 64 days before the peak, at the peak, and 64 days after the peak.

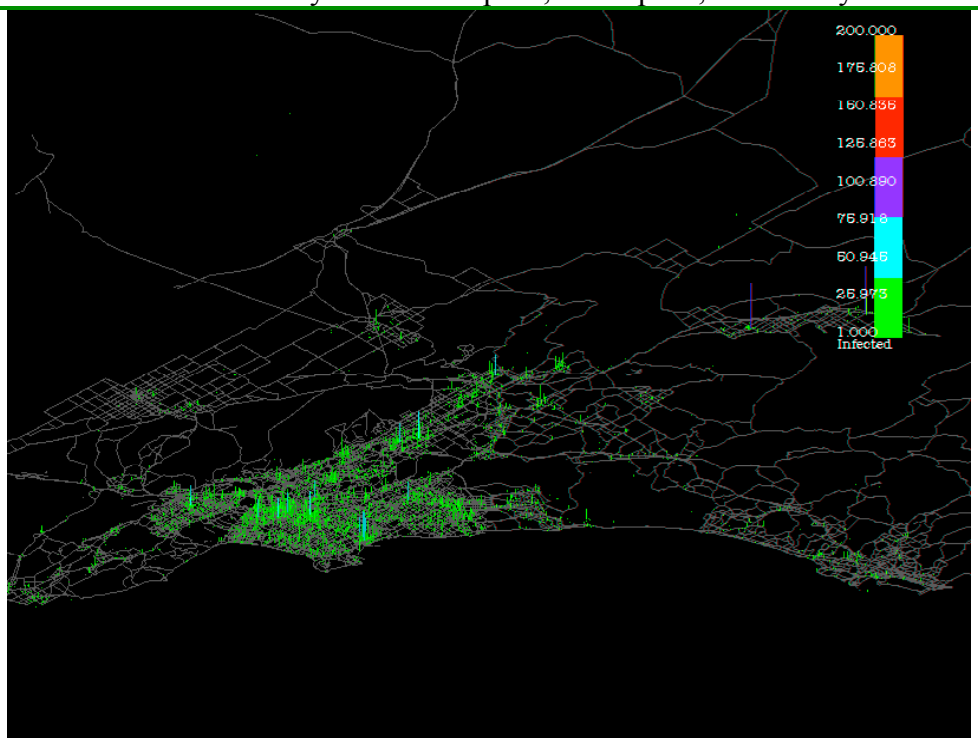


Fig. 6.3-1. The geospatial distribution of infected locations, on day 64 of the epidemic, for the base scenario, from EpiSimS simulation try30, at 10 a.m.

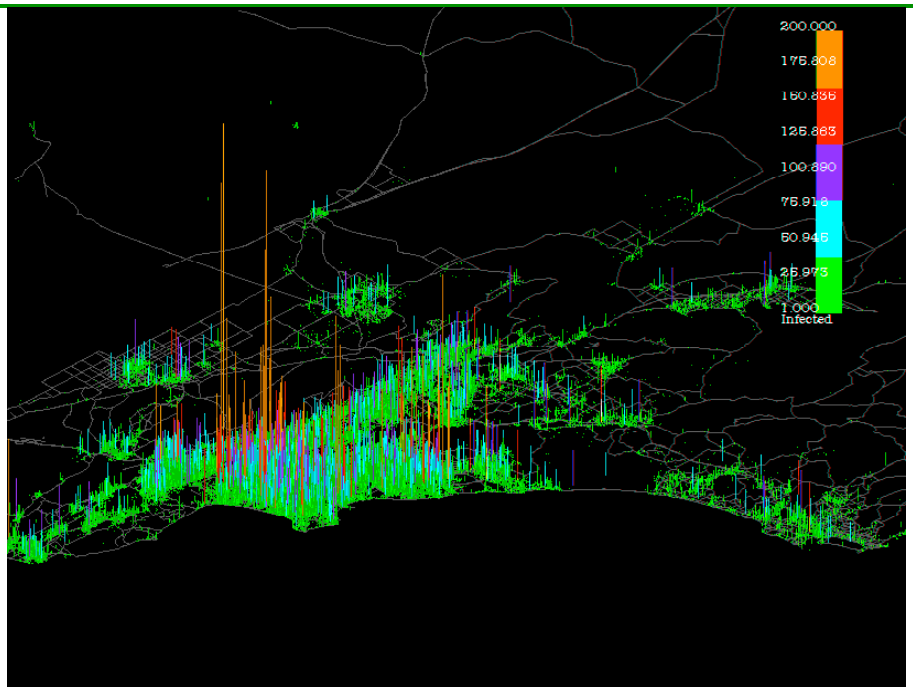


Fig. 6.3-2. The density of infected locations on day 128 of the base case epidemic, when the epidemic is at its peak, for 10 a.m.

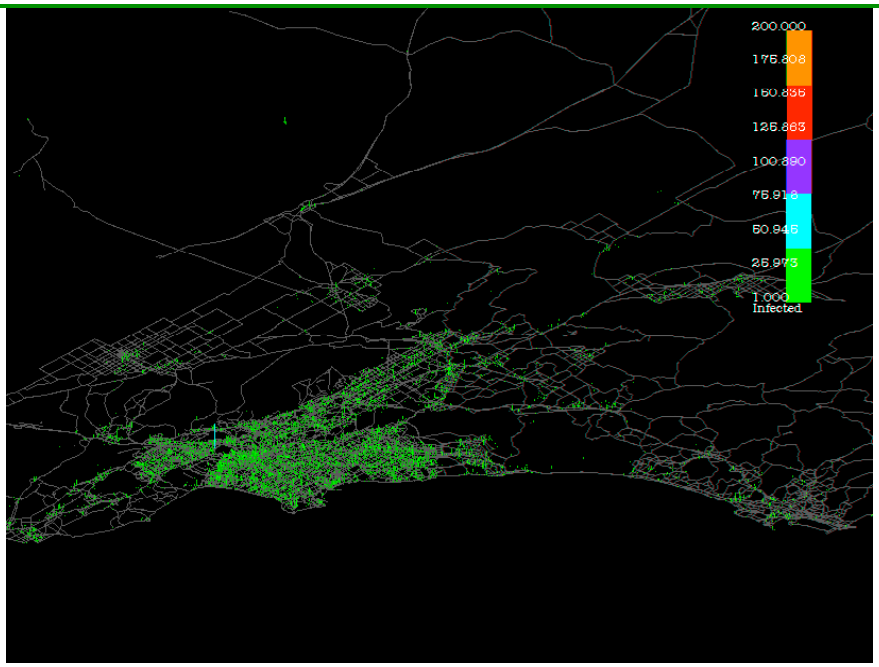


Fig. 6.3-3. The density of infected locations on day 192 of the base case epidemic, which is 64 days after the peak.

A similar view showing the terrain and cities is displayed in Fig. 6.3-4.

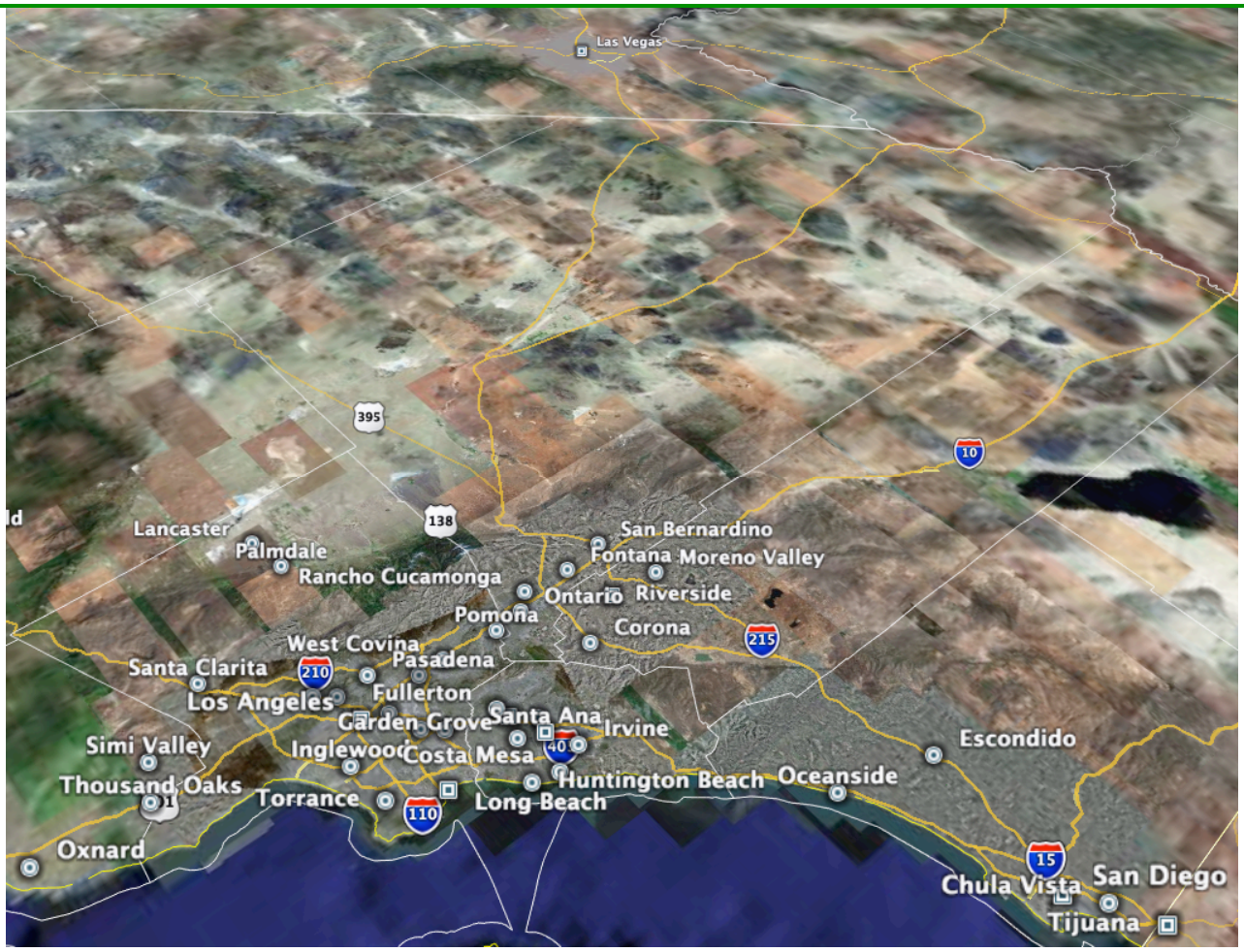


Fig. 6.3-4. Google Earth view of the five-county region, covering the same territory as Figs. 6.3-1 to 6.3-3. County boundaries are shown in white.

The epidemic simulated in EpiSimS can be aggregated at various geographical levels. The fraction of the population that is infected is shown in Fig. 6.3-5 for each of the five counties. The stochastic nature of disease transmission is evident early in the epidemic, but after one in every thousand persons are currently infected, there is enough interaction between counties that the five epidemics fall into lockstep. In the five counties, the peak incidence rates occur within two days of each other.

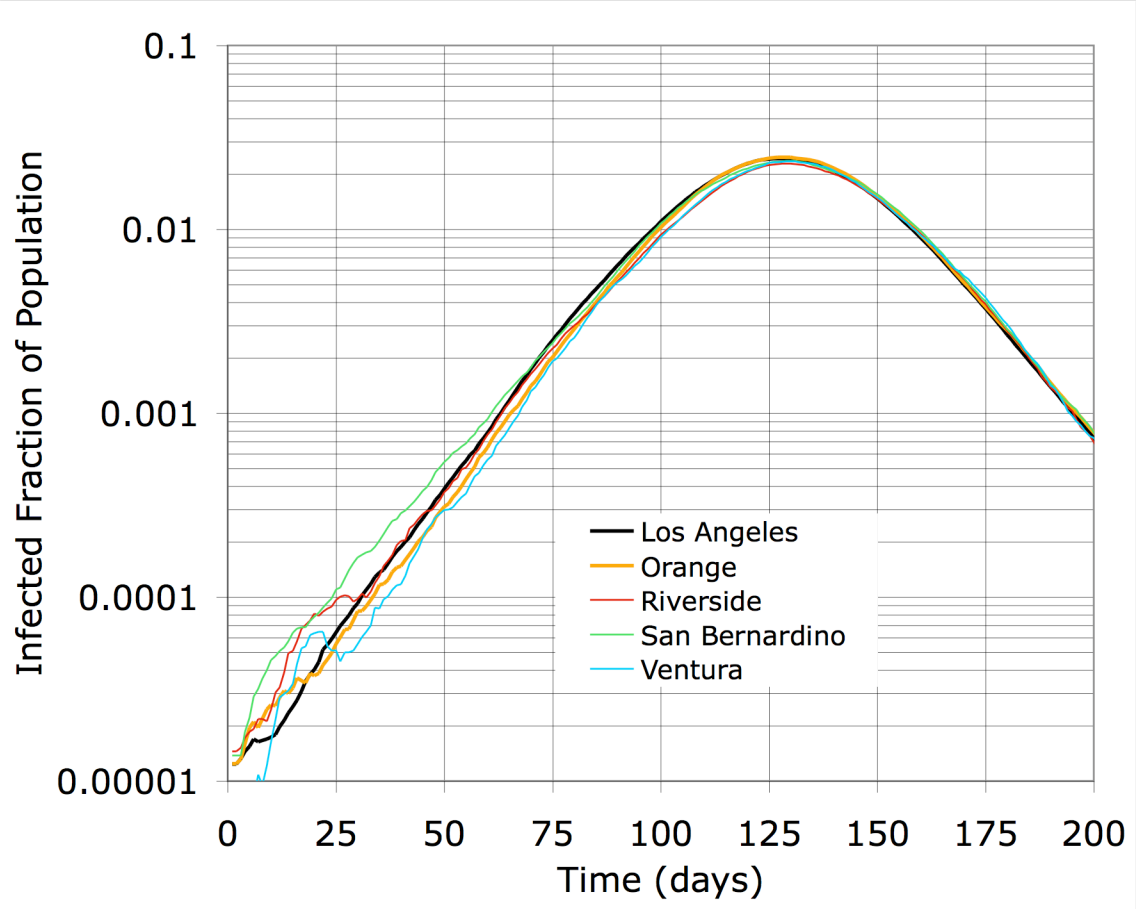


Fig. 6.3-5. The fraction of the population that is currently infected (including incubating, subclinical, and symptomatic states) as a function of time for each of the five counties.

The five counties show similar, but not identical, peak incidence rates. The perimeter counties (Riverside at 2.27%, San Bernardino at 2.36%, and Ventura at 2.34%) attain slightly lower attack rates than the central counties (Los Angeles at 2.43% and Orange at 2.47%), where these peak incidence rates specify the percentage of the population that is infected at the epidemic peak.

The overall attack rate (the fraction of the county population that ever gets infected) can also be aggregated by county. These results are shown in Table 6.3-1.

County	Population	Peak Incidence	Attack Rate	Day of Peak
Los Angeles	9,366,843	2.43%	23.9%	128
Orange	2,812,102	2.47%	23.8%	129
San Bernardino	1,672,705	2.36%	23.7%	130
Riverside	1,514,716	2.27%	22.2%	129
Ventura	740,166	2.34%	22.5%	129
Five-county	16,106,535	2.41%	23.6%	129.5

Table 6.3-1. Breakout by of the number of individuals in the synthetic population

6.4 Hospitalization Bed Surge Capacity Analysis

Currently, EpiSimS does not explicitly identify hospitals as such, but instead, they are classified as workplaces. However, hospitals are included in the overall design of the simulation, so a post-processing analysis approach was applied to determine the impact of pandemic influenza on hospital admissions. Hospital information such as location and bed count were obtained from the 2004 NGA's Emergency Response data. There are 195 hospitals, and a total of 46,809 beds within the five counties under study (Table 6.4-1). Based on the average occupancy percentage in California and average response capacity studies, we assumed that only 10% of all hospital beds would be vacant and available during the influenza pandemic [LAO 2003, Rubinson 2005, AHRQ 2005]. These beds are assumed to be licensed, physically available, and have staff on hand to attend the patient who occupies the bed. We used the critical benchmark for all states set by the Health Resources and Services Administration (HRSA), to determine surge capacity, which specifies that an additional 500 beds per 1 million population would be available in all counties (Table 6.4-1).

County	Hospitals	Beds	Surge
Los Angeles	109	30,195	11,186
Orange	35	7,826	2,944
Riverside	17	3,036	1,181
San Bernardino	26	4,355	1,651
Ventura	8	1,397	549
TOTAL	195	46,809	17,510

Table 6.4-1. Hospital beds by county. Surge includes 10% of total beds, plus 500 additional beds per 1 million population.

Hospitalization rates for the baseline scenario were obtained from two sources: 1) FluAid 2.0 software available from CDC [Meltzer 2000] and are given for different age risk groups in Table 6.4-2; and 2) The 2005 UK flu pandemic contingency plan, which states that for typical flu seasons, 20% of total cases visit a general practitioner (GP), and that 1 in 30 of these GP visits results in hospitalization.

High Risk/Non-High Risk	Minimum	Mean	Maximum
0-18 years	2.3	3.4	11.9
19-64 years	1.01	4.55	7.89
65+ years	5.5	10.75	16

Table 6.4-2. Age and risk hospitalization rates, per 1000 cases.

We assumed that infected individuals would seek medical care at their local hospital. Therefore, we obtained household IDs for all infected people and based on the hospitalization rates mentioned above, we placed them in hospitals located in the county where they live. Furthermore, we set the average length of stay in the hospital to 10 days, based on the epidemiology of influenza.

Our simulations start with 202 index cases infected with a newly emergent influenza virus. The 202 initially infected people were chosen at random from the five counties as follows: 116 from Los Angeles (0.00123% of its population), 35 from Orange (0.00124%), 22 from Riverside

(0.00145%), 23 from San Bernardino (0.00137%), and 6 from Ventura (0.00081%). We did not start with one initial infected person because given the stochasticity in the model most epidemics that start with one index case die out before becoming established.

Baseline

The number of people infected with influenza for the five counties of the base case pandemic influenza scenario, are shown in Fig. 6.4-2. The epidemic peaks around day 129 and it is almost over by day 235, at which about 25% of the population have been infected. The baseline scenario leads to 3,813,957 cases and 50,990 deaths.

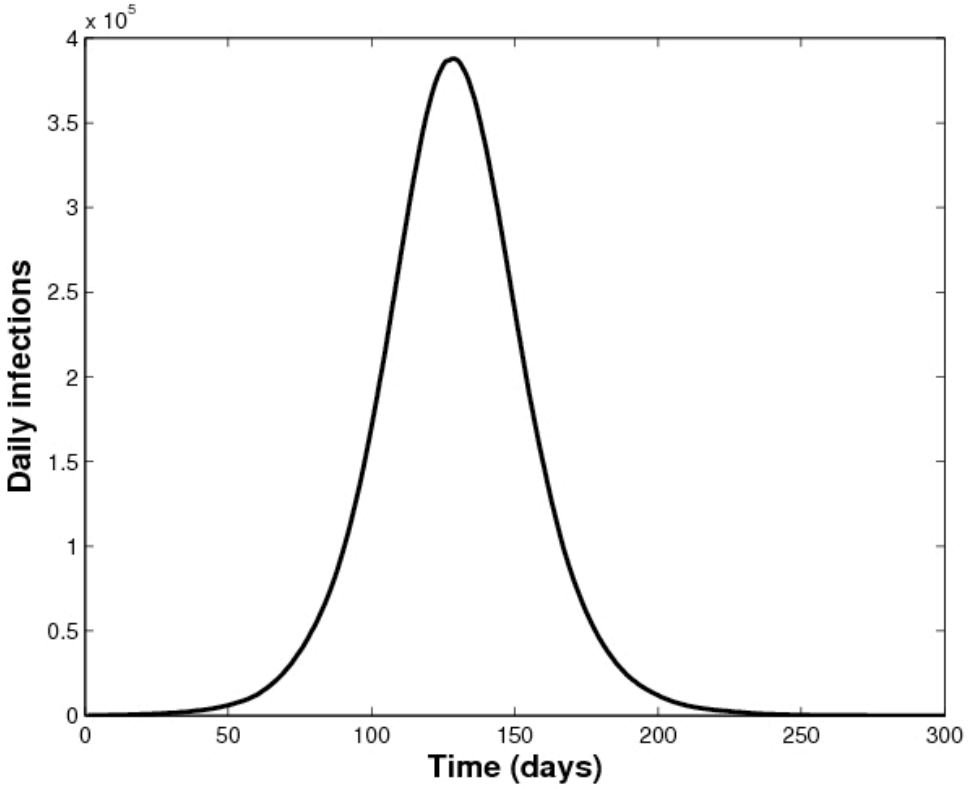


Fig. 6.4-2. Daily incidence of infections over time aggregated over all five counties (Los Angeles, Orange, Riverside, San Bernardino, and Ventura) in the absence of intervention strategies.

Table 6.4-3 shows the number and percentage of infected people by county. The disease is almost uniformly distributed among all counties, with each county having a disease prevalence of about 23% (Table 6.4-3). Even though each county started with a different number of index cases, they all eventually converged to a similar clinical attack rate. This result highlights the importance of inter-connectivity between communities/counties. Thus, it is clear that the mobility within the population plays a crucial role in the spread of a contagious disease.

County	No. of Cases	% of Cases	Total Population	% Infected
Los Angeles	2,242,332	58.8%	9,366,843	23.9%
Orange	669,818	17.6%	2,812,102	23.8%
Riverside	337,367	8.8%	1,514,719	22.2%
San Bernardino	397,682	10.4%	1,672,705	23.7%
Ventura	166,758	4.4%	740,166	22.5%
TOTAL	3,813,957	100%	16,106,535	23.6%

Table 6.4-3. Number and percentage of cases per county.

Fig. 6.4-3 shows the disease progression over time by county. We observe similar peak days and final days (when the number of cases reaches 99% of the final epidemic size) for all five counties. The epidemic peaks around day 129 in all counties and it reaches its final day around day 200. However, note that the epidemic lasts for about 300 days in all counties before it completely dies out.

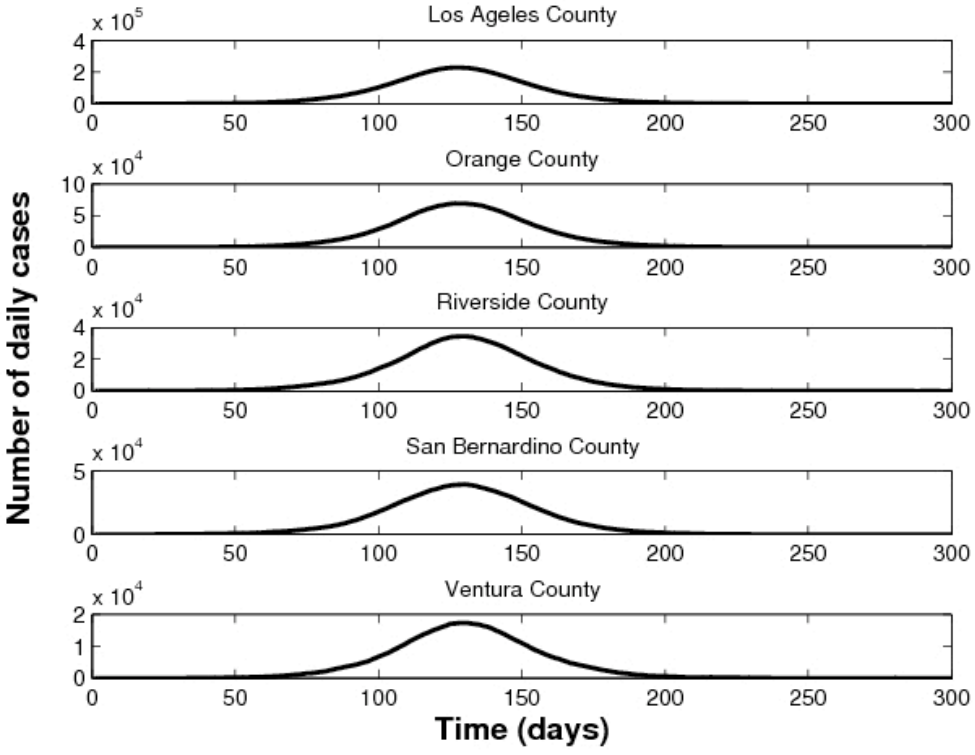


Fig. 6.4-3. New influenza infections per day by county.

The number of infected people by age group, aggregated over all counties is shown in Table 6.4-4. Our results show that adults have the highest percentage of cases; however, when the number of cases is normalized to the total population in each group, children under the age of 18 are the most affected, with 31.2% of its population being infected. This is not surprising, since children interact

with each other more than with the rest of the population and we assumed a higher infectivity rate for children. This assumption is consistent with previous pandemic influenza data, which suggests that younger people tend to be more susceptible due to lack of exposure to similar viral antigens during seasonal flu epidemics [Simonsen 1998].

Age group	No. of Cases	% of Cases	Total Population	% Infected
0-18 years	1,577,624	41.4%	5,041,103	31.2%
19-64 years	2,074,708	54.4%	9,519,517	21.7%
65+ years	161,625	4.2%	1,545,915	10.4%

Table 6.4-4. Number and percentage of cases by age group.

Fig. 6.4-4 shows the number of daily cases by age groups aggregated over all five counties. The epidemic peaks around day 125 for the youngest group, day 126 for the adult group, and day 128 for the senior group. Furthermore, we observe that at the beginning, the epidemic curves for the youngest and adult group overlap, but as the disease progresses they separate from each other. Fig. 6.4-5 shows the same data as Fig. 6.4-4, but normalized to the total number of people in each age group. We observe that the population under the age of 18 is the most affected when compared to the rest of the population (also shown in Table 6.4-4).

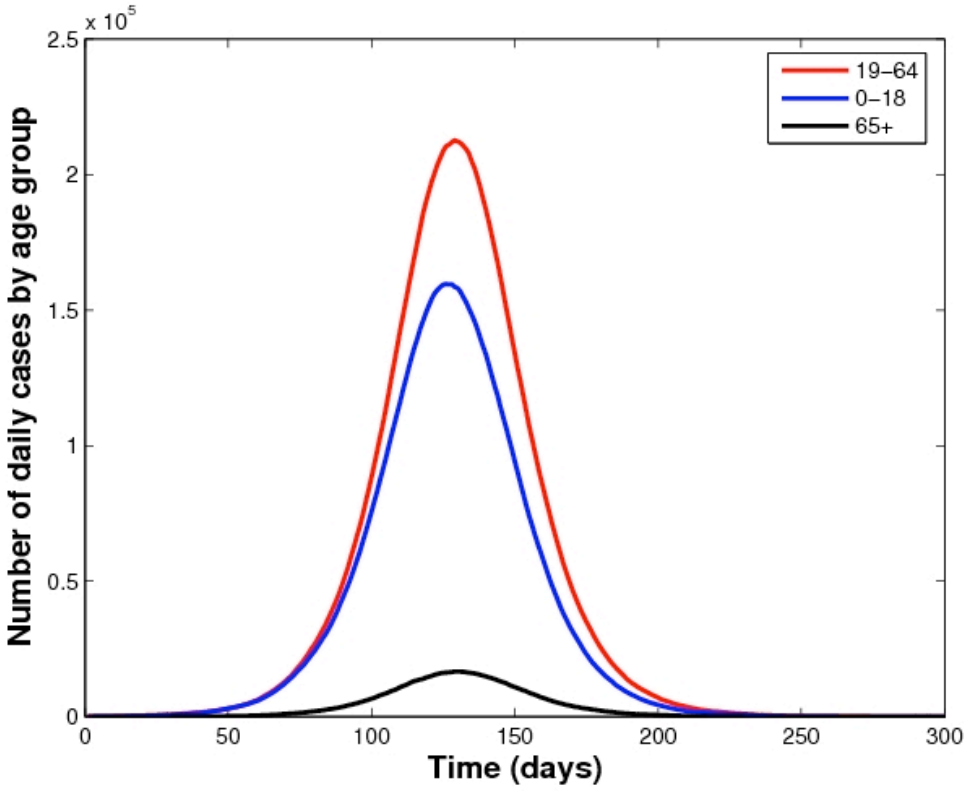


Fig. 6.4-4. New influenza infections per day stratified by age group aggregated over all five counties (Los Angeles, Orange, Riverside, San Bernardino, and Ventura).

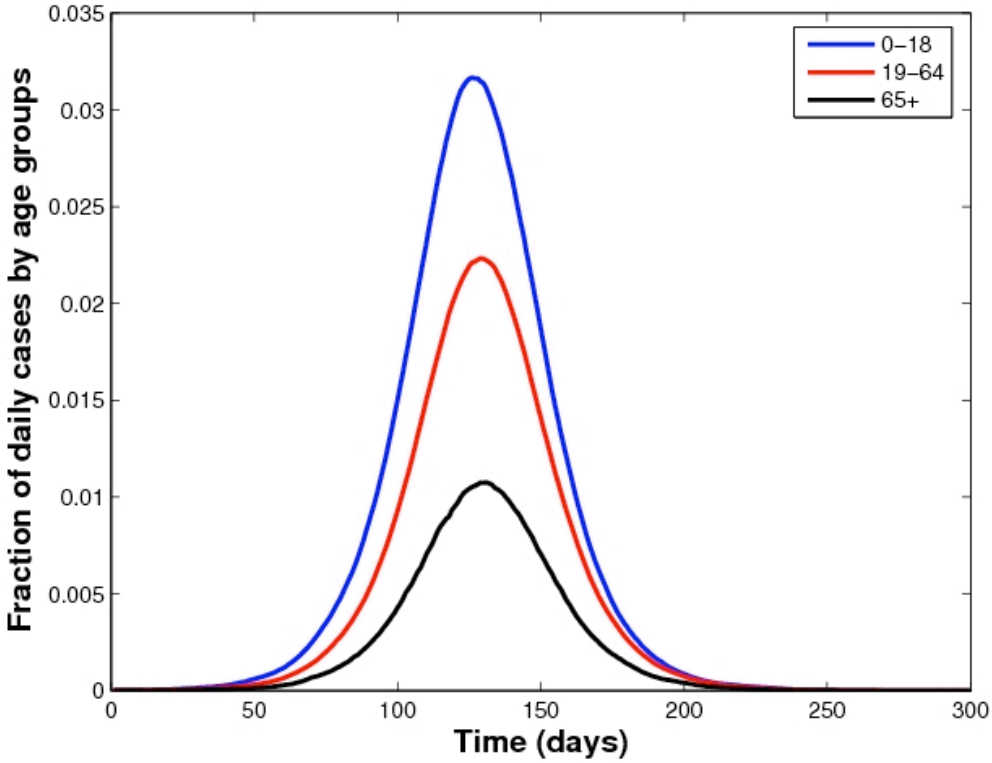


Fig. 6.4-5. Fraction of new influenza infections per day stratified by age group aggregated over all five counties. Note that the 0-18 year old age group is the most affected

Fig. 6.4-6 shows the number of daily cases stratified by age group and county. We observe different attack rates for each age group for all counties. Note that the epidemic curves for the young and elderly for Riverside and San Bernardino County have very similar attack rates. There are several possibilities that can be causing this pattern. One possibility is that the population between the ages of 0 and 18 is larger in these two counties when compared to the other three counties, and therefore they have a greater chance of becoming infected. Another possibility is that he mixing among children is higher in these two counties, resulting in more cases for these particular age groups.

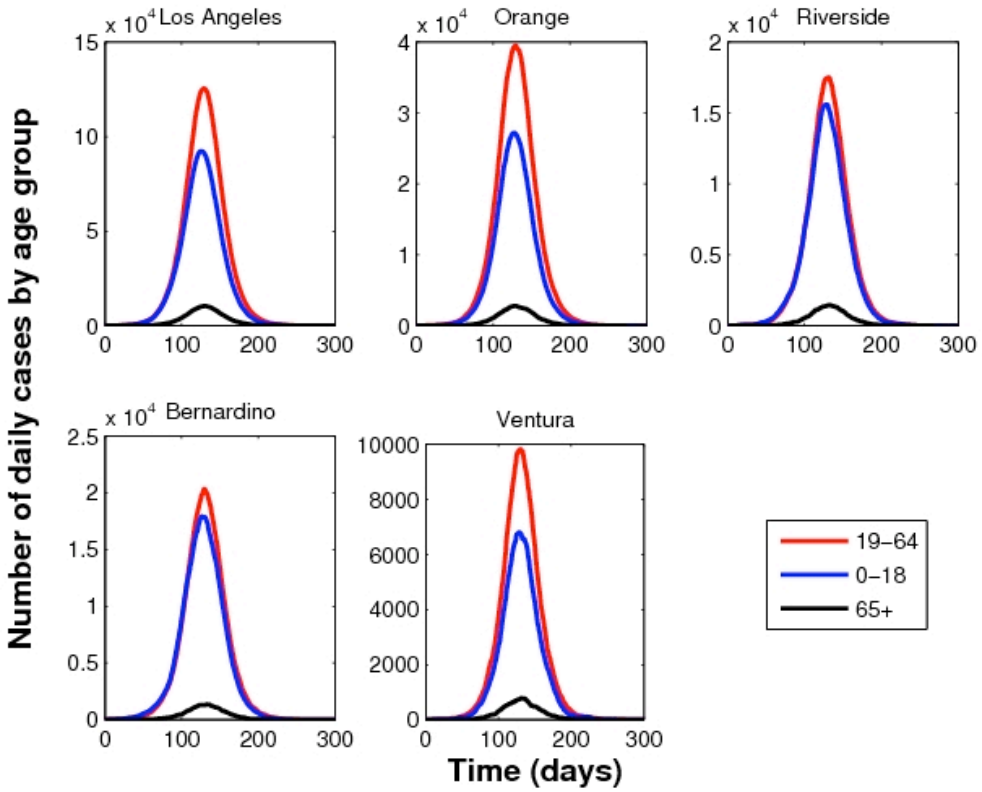


Fig. 6.4-6. New influenza infections per day stratified by age group and county.

Hospitalizations

For the baseline parameter values, our results indicate that a pandemic influenza could cause a minimum of 6,608 hospitalizations, and a maximum of 37,725 hospitalizations in all five counties (Table 6.4-5). Based on the assumption previously stated for surge capacity, an additional 3,807 and 24,992 beds could be needed if the new emergent flu virus has hospitalization rates similar to those stated by the CDC; and 12,693 additional beds if the hospitalization rates are similar to those stated by the UK contingency plan (Table 6.4-5, Fig. 6.4-7).

County	Hospitalizations			
	Minimum	Mean	UK	Maximum
Los Angeles	3,894	9,772	14,949	22,171
Orange	1,133	2,915	4,465	6,547
Riverside	601	1,446	2,249	3,389
San Bernardino	697	1,679	2,651	3,981
Ventura	283	728	1,112	1,637
TOTAL	6,608	16,540	25,426	37,725

Table 6.4-5. Number of hospitalizations by county

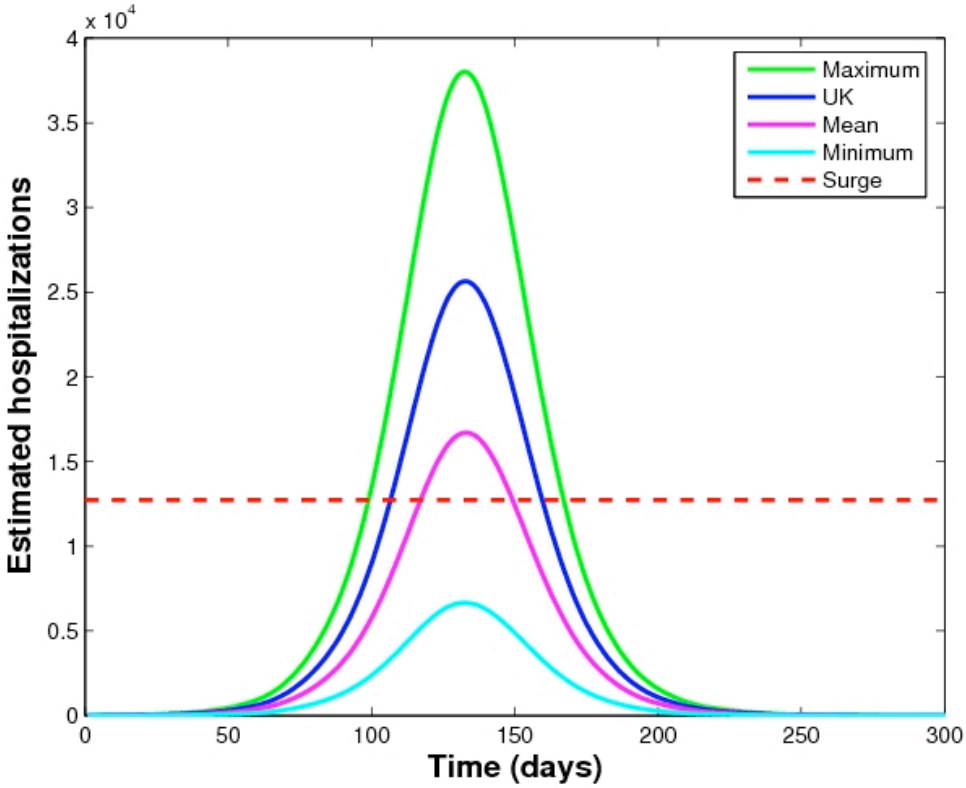


Fig. 6.4-7. Estimated number of hospitalizations aggregated over all five counties in the absence of interventions.

In Los Angeles County the total number of hospitalizations during a pandemic influenza could range between 3,894 and 22,171; resulting in 2,070, 7247, and 14,469 patients without beds for the mean, UK, and maximum hospitalization rates, respectively (Table 6.4-5, Fig. 6.4-8). In Orange county the number of hospitalizations could range between 1,133 and 6,547; and an additional 726, 2,276, and 4,358 beds could be needed for the mean, UK, and maximum hospitalization rates, respectively (Table 5, Fig. 8). In Riverside county, the number of hospitalizations could range between 601 and 3,389; and 385, 1,188, and 2,328 extra beds could be needed for the mean, UK, and maximum hospitalization rates, respectively (Table 6.4-5, Fig. 6.4-8). In San Bernardino County, the number of hospitalizations could range between 697 and 3,981; requiring 408, 1,380, and 2,710 additional beds for the mean, UK, and maximum hospitalization rates, respectively (Table 6.4-5, Fig. 6.4-8). Finally, in Ventura county, the number of hospitalizations could range between 283 and 1,637; in which 218, 602, and 1,127 extra beds could be needed for the mean, UK, and maximum hospitalization rates, respectively (Table 6.4-5, Fig. 6.4-8).

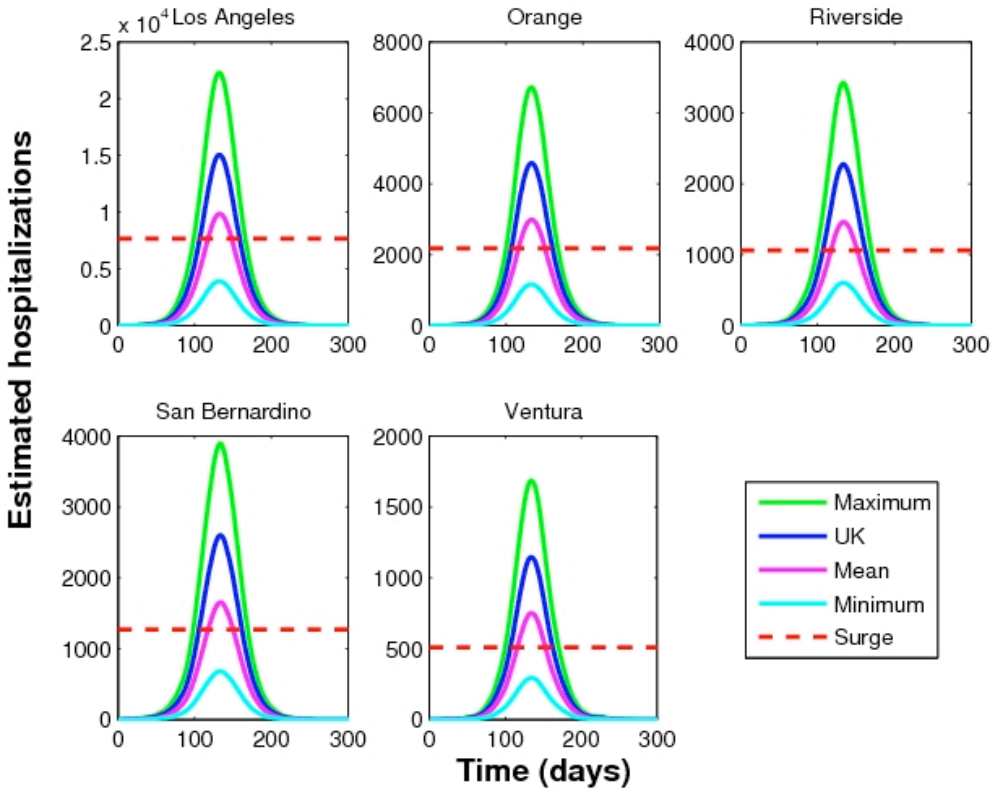


Fig. 6.4-8. Estimated number of hospitalizations and surge capacity by county.

We observe that when we aggregate the resources over these five counties, the surge capacity threshold is exceeded by about 23% for the mean hospitalization rate (Fig. 6.4-7). However, when we analyze each county separately, we observe that the surge capacity threshold for the mean hospitalization rate is exceeded by 21%, 25%, 27%, 24%, and 30% for Los Angeles, Orange, Riverside, San Bernardino, and Ventura County, respectively (Fig. 6.4-8). In fact, Los Angeles County is the only county in which the surge capacity threshold is exceeded by under 23% (Fig. 6.4-8). Therefore, analyses based on aggregated results can be misleading and may underestimate the lack of resources needed at local hospitals.

Table 6.4-6 shows the total number of hospitalizations likely to occur during a pandemic influenza stratified by age groups. Our simulations show that the young and elderly could account for a significant number of hospitalizations. Therefore, our results highlight the need for prevention efforts for both young and older people.

Age group	Minimum	Medium	Maximum
0-18 years	3,627	5,364	18,772
19-64 years	2,094	9,439	16,368
65+ years	887	1,737	2,585
TOTAL	6,608	16,540	37,725

Table 6.4-6. Number of hospitalization by age group.

Discussion

We estimate that during the length of the pandemic, a total number of hospitalizations ranging from 6,608 to 37,725 could occur. Although we observed higher hospitalization rates in the elderly, the youngest age group accounted for more than half of the total number of influenza-related hospitalizations. Therefore, these findings emphasize the importance of providing protection to both young and older people to minimize morbidity and mortality. These results are consistent with previously published studies, which highlight the need for improved prevention efforts for these age groups.

Our results show that when we combine all the resources available in Los Angeles, Orange, Riverside, San Bernardino, and Ventura County, the surge capacity threshold is exceeded by 23% for the mean hospitalization rate. However, when we analyze each county separately, we found that all counties (except for Los Angeles) exceed the surge capacity threshold by over 23% for the same hospitalization rate. Therefore, a key finding was that aggregating the total number of resources needed during an influenza pandemic, can greatly underestimate the capacity at local hospitals.

It is important to note that the analysis in this study was based on a 25% attack rate and on the hospitalization rates provided by FluAid2.0 and the UK contingency plan. However, no one knows whether the newly emergent flu virus will be similar to the previous influenza pandemics. Therefore, these results may vary depending on the virulence the new influenza virus. However, our model can be adjusted to analyze the impact of a pandemic, if real-time data is available. Nevertheless, our results are useful in providing estimates of the potential impact of the next pandemic on health care resources.

One limitation of this analysis is that it does not provide the impact of reductions on influenza-related hospitalizations due to antiviral or vaccine treatment.

The state of the health care system in California is important because Los Angeles is one of the largest cities in the nation and a major attractor for immigrants. Furthermore, studies on medical care capacity in Los Angeles have raised concerns about the hospitals' ability to respond to seasonal influenza outbreaks. Therefore, if the current health care system is not able to handle patients during seasonal influenza outbreaks, it is impossible to assume that we will be ready to manage a pandemic. Thus, preparedness is a must if we want to reduce morbidity and mortality during an influenza pandemic.

Based on this analysis of hospitalization, there are several implications for public policy. Each hospital should be investigated individually to determine the true surge capacity. Both young and old should be provided protection to minimize morbidity and mortality. Planning requires knowledge of the virulence of the newly emergent influenza virus and susceptibility of the population to the virus. Finally, there are not enough beds for the levels of hospitalizations resulting under this scenario: a significant increase in surge capacity is needed.

7 Results: Effectiveness of Disease Intervention Strategies

7.1 Case 1: 2% Anti-Virals, No Pre-Vaccination

In the antiviral treatment scenarios, courses of antiviral medicines are stockpiled, with enough courses (or regimens) to treat a specified percentage of the population. The scenarios examined are: stockpile levels to treat 2% or 4% of the population. The antivirals are given to 1) persons who present influenza symptoms, and 2) named contacts of such symptomatic persons, in particular their household members and some fraction of their co-workers or school-mates. Antiviral treatments are dispensed until either they run out, or the epidemic is stopped. Antiviral treatments are dispensed by contact tracers. The number of contact tracers, and the number of contacts they can trace per day are specified. The antiviral treatment must be received either prior to onset of symptoms, or else within 48 hours of symptom onset, to be effective.

Antiviral treatment is implemented into EpiSimS by defining distinct disease states depending on whether antiviral treatment has been received, and by adjusting the disease state transition probabilities according to whether antiviral treatment has been received. The first disease state to consider is "uninfected-susceptible". If a person receives antiviral treatment prior to exposure to an infectious dose of virus, that person is marked as having received antiviral treatment. If, subsequently, that person does receive an infectious dose in the course of his activities in the simulation, such person marked as having received antiviral treatment has a 30% chance to transition directly to an effectively immune state, and a 70% chance of transitioning to the latent-incubating disease state. Such a person will maintain the mark denoting that he has received antiviral treatment. This sequence might be expected when a person becomes symptomatic, then has all her household members treated with antivirals, and then exposes a household member who is acting in a care-taking capacity. The alternative way for a person to find himself in the latent-incubating disease state with antiviral treatment is to become infected first, and then to be named as a contact by a symptomatic person. Both routes are possible in the EpiSimS simulation.

There are four distinct disease states that a latent-incubating person undergoing antiviral treatment can transition to. These are 1) recovered-immune, 2) treated-subclinical-infectious1, 3) treated-symptomatic_non-circulating1, and 4) treated-symptomatic-circulating1. A person in the latent-incubating state will transition out of that state with a probability per unit time determined by the incubation stage sojourn time histogram described in section 3.2. When a person receiving antiviral treatment does make the transition out of the latent-incubating state, he will have a 60% chance of transitioning directly to the recovered-immune state. On the other hand, a person in the latent-incubating state not receiving treatment will have no chance of transitioning directly to the recovered-immune state, but will always go into one of the infectious states.

When an incubating person receiving antiviral treatment transitions out of the incubating state, he will have a 13.3% chance of transitioning into the treated-sub-clinical-infectious1 state. The sojourn time in this state is one day shorter than the normal infectious state (achieved by advancing the transition histogram by 24 hours). A person in a subclinical infectious state will be half as infectious as a symptomatic person, and will continue to engage in their normal activities.

The remaining 26.7% of latent-incubating-antiviral cases will transition to one of two treated-symptomatic states. There are two such states, one for persons who stay at home, and the other for those that continue to circulate even though they are symptomatic. The allocation between

circulating-symptomatic and stay-home-symptomatic states depends on the age of the person. For adults and seniors, half of symptomatic cases will stay home. For students, 75% will stay home, and 80% of pre-schoolers will stay home.

Persons in the treated infectious states transition out of those states according to the appropriate histogram. 2% of people transition into death, and 98% transition into a recovered-immune state.

Fig. 7.1-1 shows some EpiSimS simulations of the scenario with antivirals stockpiled to treat up to 2% of the population. There are 202 index cases infected at the start of the simulated epidemic.

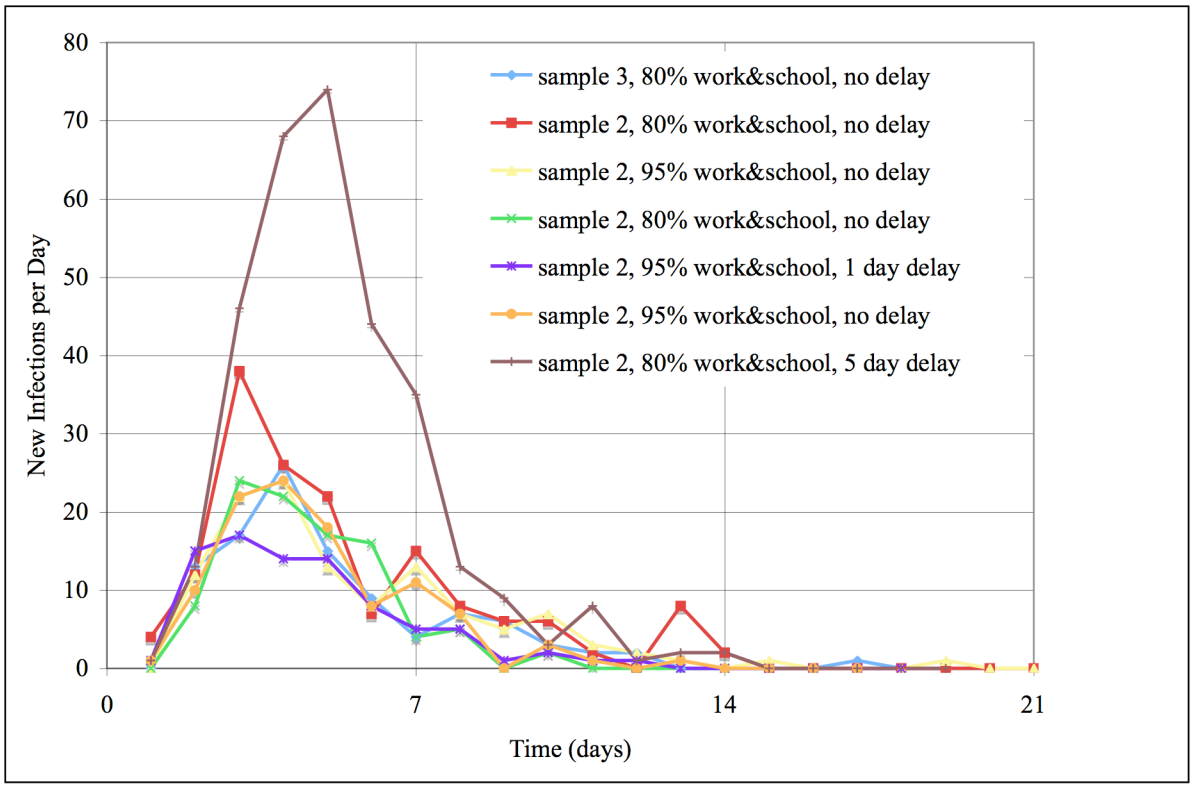


Fig. 7.1-1. New infections per day for seven EpiSimS runs with the antiviral treatment scenario, for different sets of 202 individuals infected at day 0. For some runs, the antiviral treatment was delayed to as if the disease were not immediately recognized.

In the EpiSimS results shown in Fig. 7.1-1 for antiviral treatment of contacts, all symptomatic cases were addressed by the contact tracing response. The epidemic is stopped in all the seven simulations shown. The stochastic nature of small outbreaks is clear. The epidemic curves of new cases per day as a function of time can not be reproduced with the scoping model by a simple reduction in the transmissions per day per infectious person to account for antiviral treatments. Apparently, the shortening of the effective infectious period due to antivirals is also a significant factor in the dynamics.

7.2 Case 3: 4% Anti-Virals, No Pre-Vaccination

If the antiviral treatment reduces the new-case-per-day growth rate to a negative value, the epidemic will be prevented. If the growth rate is not pushed below zero, the epidemic will eventually take off, whether antivirals are available for 2% or 4% of the population.

7.3 Case 2: No Anti-Virals, 20% Uniform Pre-Vaccination

If a vaccination can be developed for an avian influenza that jumps to humans, some people will be vaccinated before the epidemic, and some will be vaccinated during the course of the epidemic. EpiSimS can be used to simulate the administration of the vaccine to the population, accounting for resource limits in number of vaccinators, number of doses, etc. It can also implement targeted vaccination strategies, such as preferential vaccination of elderly, school-age, or some other prioritization. With this approach, a vaccination is treated as an event that happens to an individual at some time during the simulation, causing a change in the state of that individual.

Alternatively, vaccination can be viewed as a change in the population that is done prior to the epidemic. In this treatment, the population is initialized with appropriate individuals pre-vaccinated.

For a pre-vaccinated initial population, those individuals that were vaccinated will follow the following disease manifestation. Upon being exposed to an infectious dose by co-occupation of a room with an infectious person, a pre-vaccinated individual will transition immediately to a recovered-immune state with a 70% likelihood. For the remaining 30%, the vaccination was not completely effective, but their illness is milder and they are less contagious than if they had not been vaccinated. These 30% enter a vaccinated-incubating state.

These vaccinated-incubating persons will transition to a vaccinated-infectious state, with the transition following the incubation-to-symptomatic transition histogram. There are three variants of the vaccinated-infectious state: sub-clinical, circulating, and self-isolating. For all three variants, the transition from the vaccinated-infectious state is given by the histogram formulation, except that the histogram is advanced by one day. 98% of vaccinated-infectious cases transition to the recovered-immune state, and 2% of them transition to a dead state.

Besides having a day shorter average state duration, the vaccinated-infectious state differs from the (unvaccinated) infectious state by being one fifth as infectiousness. The base case infectiousness for untreated, symptomatic adults and seniors is i_0=0.005*0.0095 transmissions per contact per minute. School-agers and pre-schoolers have double the infectiousness of adults or seniors.

One third of vaccinated-infectious cases are taken to be sub-clinical. Of the symptomatic vaccinated-infectious cases, the self-isolating fraction is taken as 0.5 for adults and seniors, 0.75 for school-agers, and 0.8 for pre-schoolers. Sub-clinical infectious cases are taken to be half as infectious as symptomatic cases.

Fig. 7.3-1 shows the epidemic curve generated by EpiSimS simulation of an epidemic in the case that 20% of the population is pre-vaccinated.

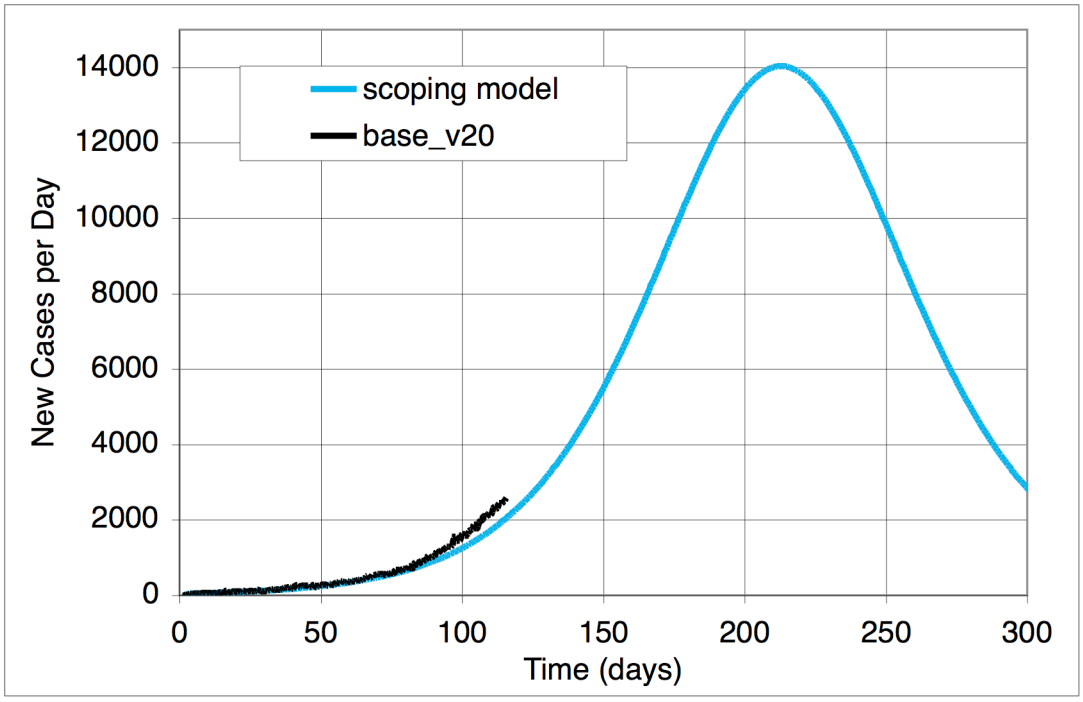


Fig. 7.3-1. The epidemic curve generated by EpiSimS simulation (run *base_v20*) of the case in which 20% of the population is pre-vaccinated prior to an epidemic. People are selected at random for vaccination, regardless of demographics. The epidemic curve predicted by the scoping model is also shown.

The scoping model can implement pre-vaccination in an approximate way by simply reducing the initial fraction of the population that is susceptible. Taking influenza vaccination to be 70% effective for prevention of disease gives that a 20% uniform vaccination rate corresponds to an initial susceptible fraction of 86% of the population. The epidemic curve (i.e. the number of new cases per day as a function of time) generated by the scoping model is shown in this case in Fig. 7.3-1 for comparison. Except for the change in the number of initially susceptible persons, the scoping model and scenario parameters are identical to those used to model the base case as described in section 2.1. Note that the scoping model does not account for the reduced infectiousness among those 30% of vaccinated persons that are not immune, of which some get sick.

For the 20% uniform pre-vaccination scenario, the scoping model gives that 10.5% of the population becomes infected during the epidemic. The epidemic reaches a peak on day 210, at which time 0.087% of the initial population is becoming infected per day.

7.4 Case 4: No Anti-Virals, 40% Uniform Pre-Vaccination

At a 40% vaccination level, the basic reproductive number drops to about 1 average transmission per case. From the base case reproductive number of 1.34, if the vaccination is 70% effective in preventing infection, the reproductive number would drop to 1.34(1-0.2*0.7)=0.9648. Fig. 7.4-1 shows the epidemic curve from EpiSimS run base_v40, which treats the scenario in which 40% of

the population, selected independently of demographics, is vaccinated. In this regime, stochastic effects are pronounced.

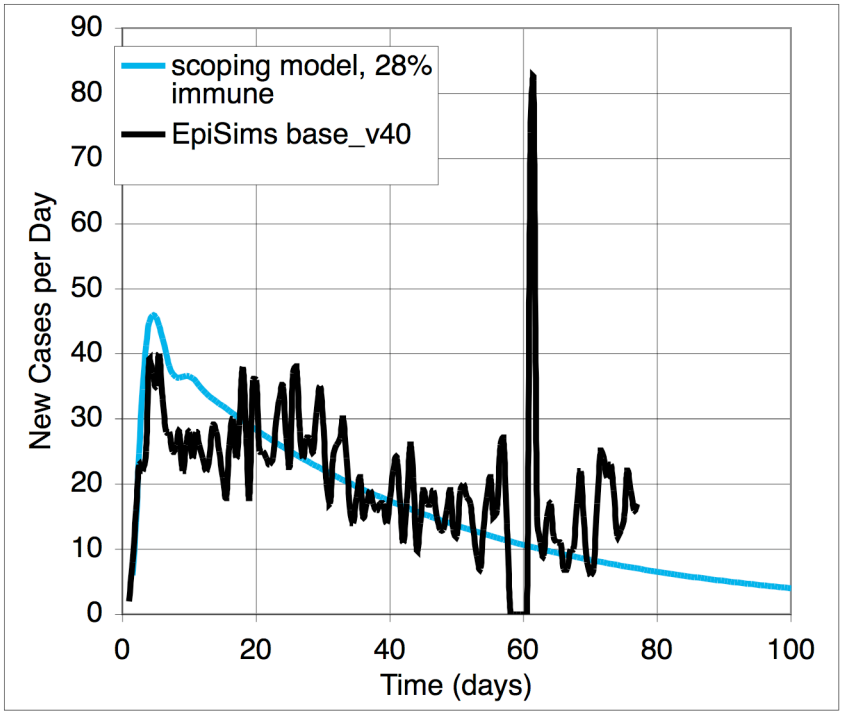


Fig. 7.4-1. The epidemic curve giving the number of new cases per day, for the scenario having 40% of the population vaccinated prior to the onset of the epidemic.

The half-day-bucket-transition model with power-law transmission scaling was used to generate an epidemic curve, which is also shown in Fig. 7.4-1. The herd immunity was set to 28%, to account for 40% of the population being vaccinated, with the vaccine having a 70% effectiveness in preventing disease. This gives a starting population of which only 72% are susceptible. Also, the basic reproductive number was reduced from 1.34 to 1.25 to provide a reasonable fit to the EpiSimS simulation of the 40% vaccination scenario. This apparently accounts for the reduced infectiousness that occurs in people that get sick but that had been vaccinated.

7.5 Case 5: No Anti-Virals, 20% Targeted Pre-Vaccination

In targeted vaccination, doses of vaccine are available for a specified fraction of the population. A program of vaccination is initiated at the start of the simulation, at simulation time = 0. The number of people that can be vaccinated per day is limited by the number of available nurses or other health care personnel, and the time it takes them to move from one location to the next, and the time spent at each location dispensing the vaccine. For targeted vaccination, we implement a prioritization in the vaccine recipients that is based on 1) the demographic groups that received high priority by the CDC guidelines during the vaccine shortage in the US during the 2004-2005 epidemic influenza season, and 2) the prioritization guidelines of the United Kingdom Health

Department. This scenario of mass-vaccination of ~20% of the population, targeted at persons aged 5 and under, and at persons 65 and older, is was constructed for consistency with the 2004-2005 flu season.

Fig. 7.5-1 shows the epidemic curve (new cases per day, versus time) for the scenario that 200 randomly selected persons are infected at time 0, and a targeted mass-vaccination program is initiated at the same time. The number of vaccinators is set to a high enough value that everyone on the target list is vaccinated early in the first day of the simulated outbreak.

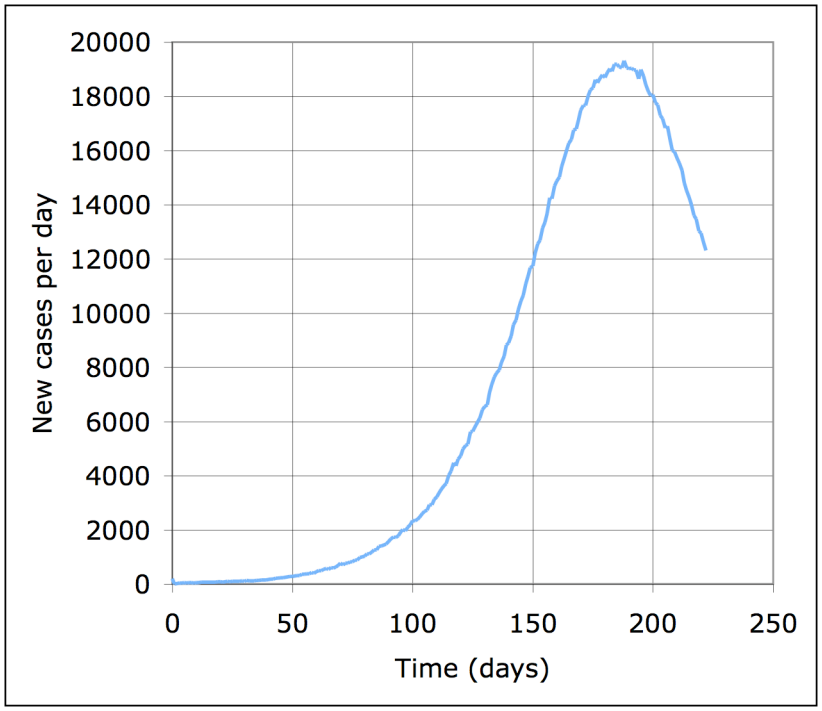


Fig. 7.5-1. The epidemic curve for a targeted mass-vaccination scenario, in which 20% of the total population receives vaccine, and children (5 and under) and seniors (65+) have priority.

An approximate surrogate for modeling the impact of mass vaccination is to simply move some fraction of the initial susceptible population to an immune state at the beginning of the simulation. As we assume that 70% of susceptible persons receiving the vaccination prior to exposure would become immune, this approximate approach can be implemented by making 70% of children age 5 and under, and 70% of seniors age 65 and older, to be immune at the start of the simulation (seniors and preschoolers combine to form roughly 20% of the population). In Fig. 7.5-2, the epidemic curve is shown for this approximation, as is the epidemic curve obtained with the complete vaccine-treated disease manifestation.

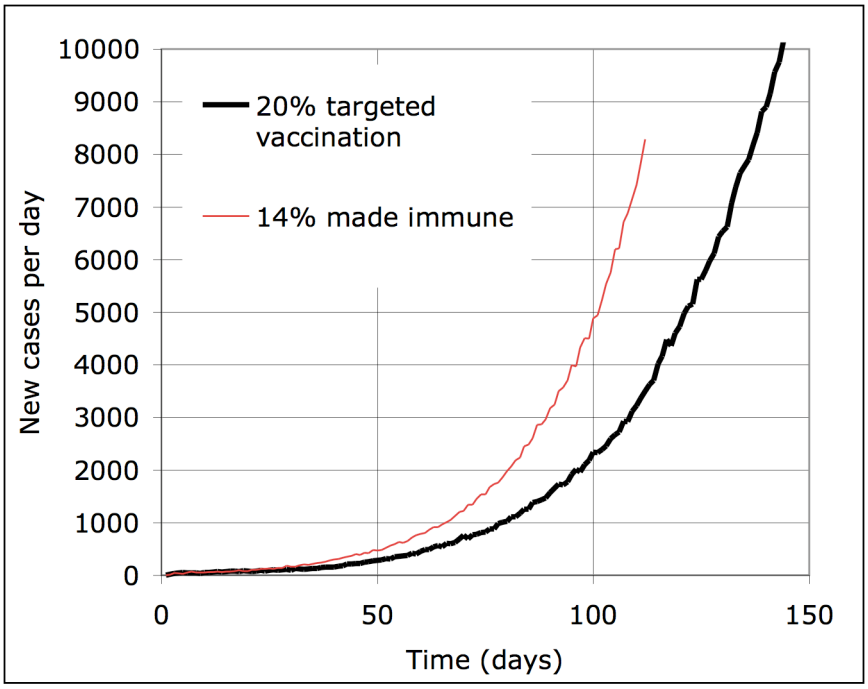


Fig. 7.5-2. Epidemic curves from 1) simple vaccine surrogate model in which 14% of the population (all selected from preschoolers and seniors is made immune) are made immune at the start of the simulation, and 2) the complete vaccine EpiSimS implementation, in which 20% of the population is vaccinated at the start of the simulation (targeting preschoolers and seniors), with proper treatment of reduced infectiousness and duration of infectiousness among vaccinated persons that do become infected.

The simplified surrogate treatment of mass vaccination by moving people directly to an immune state if their vaccine is effective is found to give about twice the epidemic growth rate as the more detailed treatment that includes the reduction in the infectious period by one day, and reduction in the infectiousness of vaccinated but infected persons.

7.6 Case 6: Masking

Another consequence mitigation strategy is to get people to wear masks. Even in the absence of effective antiviral medication or vaccine, if 50% of persons over the age of 5 wear N95 masks while they are in contact with other people, the epidemic can be prevented. Fig. 7.6-1 shows the number of new infections per day and the current number of symptomatic persons for when the masking response is used. The average number of transmissions per infected person drops to a value of 0.9, which is insufficient to maintain the chains of infection. Starting with 202 infections, there are a total of 6 fatalities. There are a total of 1253 infections, counting the 202 index cases.

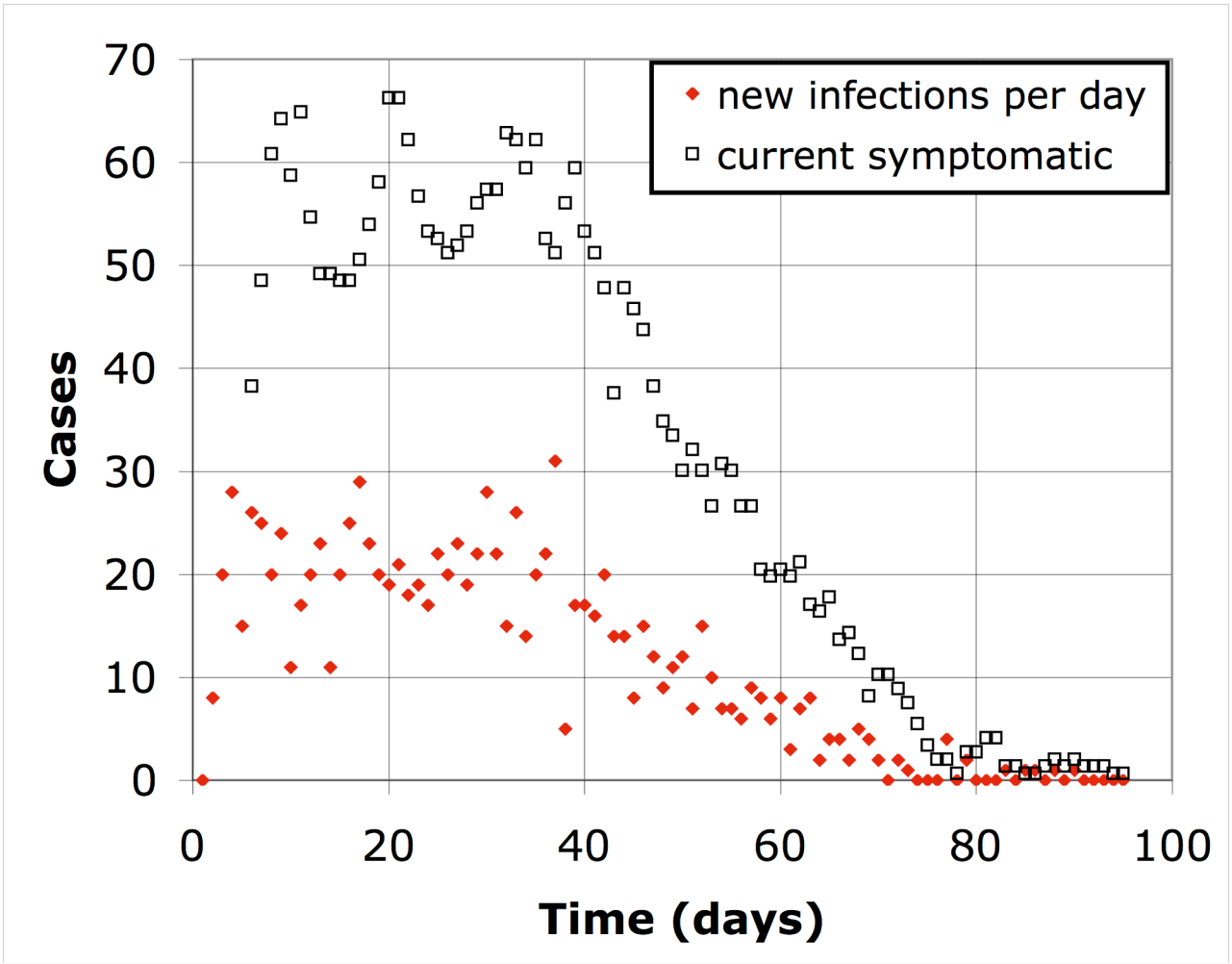


Fig. 7.6-1. The epidemic curves for a pandemic influenza in Los Angeles, for a scenario matching the base case, except that 50% of susceptible persons over the age of 5 wear ventilated N95 masks when in contact with other persons.

A second masking scenario specifies that 25% of susceptible persons over the age of five wear ventilated N95 masks when in contact with other persons. An EpiSimS simulation run for 200 days of simulation time finds that the epidemic is slowly growing, with a growth rate of 1.7% per day. The scoping model finds that this corresponds to an average reproductive number of 1.08. The epidemic curve generated by EpiSimS, and the projection constructed with the scoping model are shown in Fig. 7.6-2. The projected outbreak reaches a peak of 5008 new infections per day. The peak is attained about 300 days after reaching a new infection rate of 100 new cases per day. Stochastic effects can be clearly observed early in the epidemic. By the end of the outbreak, there will have been a total of 1,111,279 infections.

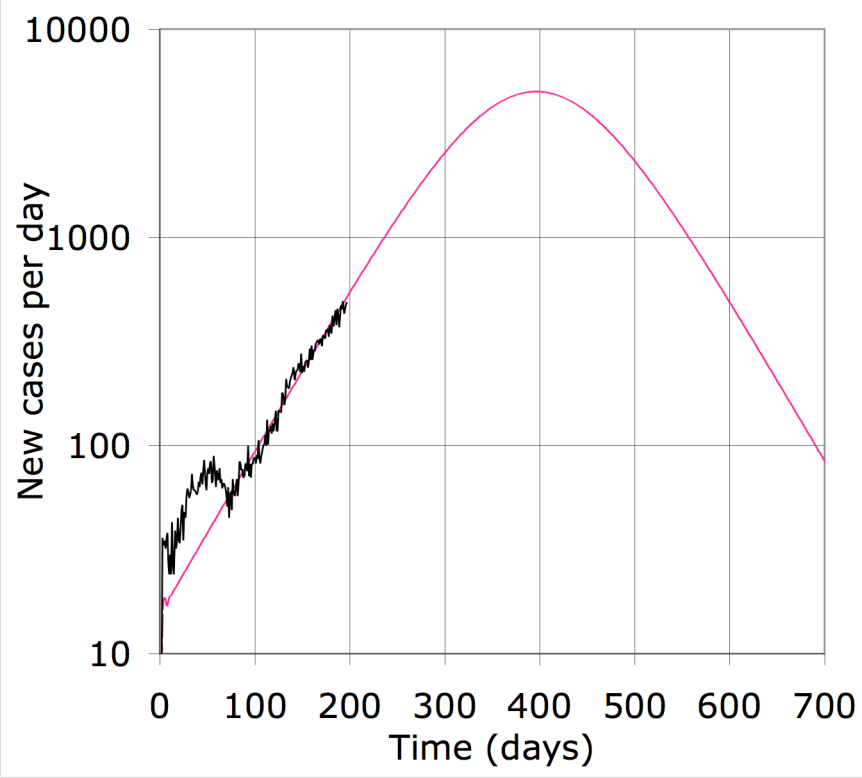


Fig. 7.6.2. The epidemic curve generated by EpiSimS for a scenario matching the base case, except that 25% of susceptible persons over the age of five wear ventilated N95 masks when in contact with other persons. The projection constructed with the scoping model is also shown as the smooth red line.

8 Known Issues and Future Work

Improved Validation and Statistics for UPMoST Information

Links as well as locations were made available in the UPMoST Household and Activity information. A link represents an aggregated set of locations. A home location and link were available for households and locations and links were available for activities. The UPMoST software required a modification to be able to read these links correctly. The link was stored as a float when read, losing digits when converted to an integer. This was changed to a double.

The UPMoST Activity information had missing and extra data. Some people did not start their activities at 00:00:00 and some continued to have activities after midnight. EpiSimS requires a 24-hour schedule for each person. Schedule generation was modified to correct these situations after multiple iterations of bad schedules. Validation of the UPMoST information would have saved us time and effort in the preprocessing.

Statistics should also be made available about the UPMoST person, household, and activity information. Knowing the distribution of the demographics over the people and households, and activities throughout the day could help us in designing scenarios and understanding results.

Checkpoint and Restart Capability

A checkpoint and restart capability is desirable for long runs that cannot be completed as a single job. Each person's health state and current schedule item, as well as any scenario commands currently in affect (ex. self-isolation, mass treatment, ring delivery) need to be written to a checkpoint file. The time interval for checkpoint files should be selectable by the user.

Regression Testing Suite

A set of tests, under version control, representing the common scenarios (ex. none, self-isolation, mass treatment, ring delivery, etc) should be developed and run for EpiSimS after major changes and when porting to other parallel architectures. These tests should have a 1K-, 10K-, 100K-, and 1M-person configuration associated with each.

Hospitals and Morgues

Hospitals should be modeled specifically in EpiSimS with attention to details such as the number of beds available and the people that actually work there (doctors, nurses, lab technicians, etc). Hospital workers will need to be marked separately from other people. This will allow for the treatment of hospital workers before others and the creation of hospital worker disease manifestations. Additional scenario commands may be required to cause people to go to a hospital based on a disease attribute (e.g. when symptoms > 3).

Because several historical smallpox cases were acquired in morgues, Morgue might be added as a new type of location. Once a person has died they can be moved to a morgue. This is a low priority, as the impact of such transmissions would be small.

New Output Events and Selection

Additional output events that would be useful are events for when a room changed disease state, and when a room was treated. Example information for each type follows.

Person infected event:

```
I <time> <person-id> <room-id> <location-id> <activity-id>
```

Room changed disease state event:

```
RD <time> <room-id> <old-disease-state> <new-disease-state> <location-id>
```

Room treated event:

```
RT <time> <room-id> <location-id> <treatment-id>

The output event types desired by the user should be made selectable as a
parameter in the configuration.
```

Enhance Treatment Delivery Scenarios

The scenario commands for mass treatment and ring treatment should be extended to include a time to stop delivery of treatments, allow treatment to individuals based on other demographics besides age, and allow selection of activity (ex. home, work, shop, school, college) that should be targeted.

More Behavior Modification Disease Intervention Strategies

Currently, the EpiSimS simulation supports a couple of behavior modifications, self-isolation and the use of masks. We would like to add more behavior modifications, such as the use of gloves, hand hygiene, respiratory cough etiquette, limiting travel, working from home, staying home from school, and closures by activity. They should be added in a general manner to allow easy addition of new behavior modification disease intervention strategies.

Multiple Diseases or Mutating Disease Simulation

Explore how multiple diseases (ex. Flu and Pneumonia) and mutating diseases (ex. flu) can be modeled simultaneously. Research how disease manifestations combine in individuals. Design and develop the enhancements required.

Partitioning Enhancements

We are in the process of testing the sublocation model using a maximum count of the number of persons at a location for specified time intervals during the day. We intend to compare this to the results given with what we currently do: use the total number of people at a location during the day. The idea is to gather statistics and see if there is a change because of more sparsely populated rooms in our current model.

Locations are currently assigned to the processors in a run randomly. The current load distribution may not be balanced. Exploring the partitioning of locations based on activity or geographical proximity could make a difference in the efficiency of the code's message passing usage.

Set-up Tool

Perfect a Set-up Tool that leads the user through a step-by-step process of simulation set-up. The tool will check for consistency between the files and generate the configuration file needed to run the simulation. The tool should provide documentation of the simulation run and enforce a consistent structure to the input and output files. The tool is linked to other pre-processing programs: InitializeHealth, BuildDiseaseManifestation, and ScenarioBuilder.

Lower Resolution Synthetic Population

In order to be able to simulate larger and larger populations or run lower fidelity EpiSimS models, we need to aggregate data or eliminate data. One approach is to use "tract" as the location resolution to aggregate locations. Schedule and partition files would be generated using these "tract" locations. Then 99% of the households would be eliminated randomly. This would also remove the people associated with those households. We will explore how the epidemic results of this smaller scale model compare to the corresponding high-fidelity version.

Spatial Dynamics of Epidemics

Currently, to assemble the data associated with the spatial dynamics of an epidemic over time requires post-processing of the event output along with data from the entities, schedules, etc. This can be time-consuming and tedious. Another approach is to collect counts of the epidemic parameters (# infected, infectious, recovered, dead, etc.) per location as the simulation progresses. Counts would be modified as people arrive/leave a location and when people change disease state. These counts could be reported per day or at a user selected time interval. Counts over the entire

simulation could be summed up across all locations on a processor and gathered and reported by the MASTER node. This would also be useful in watching the progress of a simulation as it is running.

New Architectures

EpiSimS will be ported to other LANL Institutional Computing clusters such as TLC and Coyote with architectures (ex. processors, memory size, interconnect, file system) different from Pink and our local cluster (delibes). Faster processors, more memory, faster interconnects, and parallel file systems will allow us to run larger population models. Some design changes, message passing tuning, and use of parallel I/O software will be required.

9 References

Albert R, Barabasi AL, *Statistical Mechanics of Complex Networks*, Rev. Modern Physics, Vol. 74, pp. 47-97, Jan 2002.

The Agency for Healthcare Research and Quality (AHRQ). Optimizing Surge Capacity: Hospital Assessment and Planning. 2004. Available from: URL: http://www.ahrq.gov/news/ulp/btbriefs/btbrief3.pdf.

Barabasi AL, Bonabeau E, *Scale-Free Networks*, Scientific American, May 2003.

Barrett CL, Eubank SG, Smith JP, *If Smallpox Strikes Portland*, Scientific American, March 2005.

Barry JM, *The Great Influenza: the Epic Story of the Deadliest Plague in History*, Viking Press, 2004.

Blagoev K, Germann TC, Johnson NL, Kadau K, Lomdahl PS, Macken CA, McPherson TM, *Large-Scale Epidemiological Simulation: Building New Capabilities on Renowned Science Foundations*, Los Alamos National Laboratory technical report LAUR-043867, 2004.

DH/HPIH and SD-Immunisation Policy, Monitoring and Surveillance. Pandemic Flu: UK Influenza Pandemic Contingency Plan. 2005.

Elder AG, O'Donnell B, McCruden EAB, Symington IS, Carmen WF., *Incidence and recall of influenza in a cohort of Glasgow healthcare workers during the 1993–94 epidemic: results of serum testing and questionnaire*. Br Med J, Vol. 313, pp. 1241–1242, 1996.

Eubank S, Goclu H, Kumar A, Marathe M, Srinivasan A, Totoczkal Z, Wang N, *Modelling disease outbreaks in realistic urban social networks*, Nature, 2004, 429, pp. 180-184.

Hyman JM, LaForce T, *Modeling the Spread of Influenza among Cities*, Chapter 10 in Bioterrorism: Mathematical Modeling Applications in Homeland Security, HT Banks and C Castillo-Chavez, eds., SIAM series: Frontiers in Applied Mathematics 28 (2003).

Halm, *Management of Community-Acquired Pneumonia*, NEJM, Vol. 347, pp. 2039:2045, Dec 19, 2002.

Legislative Analyst's Office (LAO). Health & Social Services. 2003. Available from: URL: http://www.lao.ca.gov/analysis _ 2002/health _ ss/healthss _ anl02.pdf.

Longini IM, Halloran ME, *Strategy for Distribution of Influenza Vaccine to High-Risk Groups and Children*, Am. J. of Epidemiology, 161,4,Feb 15, 2005.

Longini IM, et.al. *Containing Pandemic Influenza with Antiviral Agents*, Am. J. of Epidemiology, Vol. 159, No. 7, April 1, 2004.

McCullers JA, Rehg JE, *Lethal Synergism between Influenza Virus and Streptococcus pneumoniae: Characterization of a Mouse Model and the Role of Platelet-Activating Factor Receptor*, J. of Infectious Disease, Vol. 186, pp. 341-50,2002.

Martin I. Meltzer, H.A. Shoemake, M. Kownaski, and R. Crosby. Fluaid 2.0. 2000. FluAid 2.0, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Lewis Rubinson, Jennifer B. Nuzzo, Danile S. Talmor, Tara O'Toole, Bradley R. Kramer, Thomas V. Inglesby, and the Working Group on Emergency Mass Critical Care. Augmentation of Hospital Critical Care Capacity After Bioterrorist Attacks or Epidemics: Recommendations of the Workng Group on Emergency Mass Critical Care. *Crit. Care Med.*, 33(10):E1–E13, 2005.

Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, Fukuda K, *Influenza-associated hospitalizations in the United States*, JAMA, Vol 292, No. 11, pp. 1333-40, Sept 15, 2004.

UK Health Deptartment*, Influenza pandemic contingency plan,* (www.dh.gov.uk) Mar 2005.

US FDA center for drug evaluation and research, http://www.fda.gov/ cder/ drug/ antivirals/ influenza/ default.htm

Paule van den Driessche and James Watmough. Reproduction Numbers and Sub-Threshold Endemic Equilibria for Compartmental Models of Disease Transmission. *Math. Biosc.*, 180(1-2):29–48, November-December 2002.